Essays on Credit Risk

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Philosophy in Banking & Finance

presented by
Marc Arnold
from Zurich, Switzerland

approved in October 2011 at the request of
Prof. Dr. Alexander Wagner
Prof. Dr. Kjell Nyborg

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorizes the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, October 26, 2011

Chairman of the Doctoral Committee: Prof. Dr. Dieter Pfaff

Table of Contents

List of Figures

List of Tables

PART I: INTRODUCTION

1. Credit Risk and Credit Derivatives

One of the main risks of granting a loan or investing in a debt security is credit risk. Credit risk is the risk that the borrower of the lending contract cannot fulfill key financial obligations, such as repaying the notional or making interest payments. Credit derivatives have been developed in response to the demand by financial institutions to hedge or diversify credit risk, to take on credit exposures at lower costs, and to manage or trade credit risk independently of the ownership of the underlying asset. Most credit derivatives take the form of a credit default swap. This contract transfers the default risk of the underlying corporation or sovereign entity from one party to another against a predetermined fixed fee. Today, many varieties of this basic contract are traded in financial markets.

Credit derivatives are one of the most important financial innovation of the last decade. In a relatively short time, they have grown to become a large segment of financial markets. From 2001 to 2009, the outstanding notional amount increased from an estimated $919 billion to $30 trillion (ISDA, 2010). As a result, credit risk has gradually changed from an initially illiquid risk to one that is traded much like other fundamental factors of financial risk such as equity, currency, or interest rate risk.

The liquid trading of credit risk offers many opportunities for financial institutions and companies. It helps them to manage the credit risk exposure by allowing to insure against a deteriorating credit quality of their borrowers. For financial firms, managing concentration risk is of particular interest. The standard treatment has been to screen new loan applicants regarding the degree of additional concentration risk an approval could impose on the current lending portfolio (Batten and Hogan, 2002). Credit derivatives have the potential to implement a much more flexible approach to managing concentration risk. Large exposures to concentrated borrowers can simply be replaced with smaller and more diversified exposures. Another potential benefit of credit derivatives is their use in the management of liquidity risk. By transferring credit risk to a counterparty, the original lender may be able to substitute cash for an illiquid asset, or to conserve costly regulatory capital. The frequent trading of credit derivatives also allows lenders, borrowers, and investors to directly compare prices of bonds or loans with market prices of credit risk. Hence, credit derivatives add transparency to the pricing of credit risky claims by offering a benchmark.

The recent financial crisis, however, painfully reveals that the liquid trading of credit risk brings about not only opportunities but also challenges. Academics and practitioners argue that credit risk transfer may induce problems related to asymmetric information. For example, incentives of lenders to analyze and monitor credit quality could be reduced if they have the ability to hedge their credit exposure. Lower credit discipline could then decrease overall credit quality. While asymmetric information problems exist in most markets, the credit derivatives market is particularly vulnerable. The reason is that lending firms have a much closer relationship to their borrowers than outside investors. Hence, by the nature of the credit business, most of the major players in this market are insiders.

Furthermore, during the recent crisis, many commentators have raised the concern that while credit derivatives may allow for risk reduction at the individual entity level, it is not clear how they affect the aggregate risk in the economy at the systemic level. Instefjord (2005), for example, shows that credit derivatives enhance risk sharing but, at the same time, also lead to further acquisition of risk which eventually destabilizes the banking sector. Similarly, Duffee and Zhou (2001) argue that the introduction of credit risk transfer can cause the markets for other loan risk-sharing to break down.

The adequate pricing of credit risk is another unresolved matter. Even though this field has witnessed a tremendous acceleration in research effort aiming at a more profound understanding, modeling and pricing of credit risk, many problems are still discussed. For example, observed credit spreads are high relative to the empirical default rates and recovery rates, known as the credit spreads puzzle (Elton, Gruber, Agrawal, and Mann, 2001; Huang and Huang, 2003; Chen, 2010). Moreover, relatively little is known about how cross-sectional differences in firm characteristics affect credit risk.

From an academic viewpoint, the main objective of research on credit risk must be to contribute to the discussions of the above mentioned open challenges to support the evaluation of adequate solutions. The present dissertation follows this avenue by rigorously analyzing three important open issues of credit risk.

I am grateful to my supervisor Alexander Wagner for his great support and helpful comments on all three papers which have led to essential improvements. During the process of writing, I have also profited from valuable discussions and remarks by, among others, Rajna Gibson, Kjell Nyborg, and Ramona Westermann.

2. Summary of Research Results

This dissertation contains three independent research papers which are briefly outlined in this section.

The first paper "Private Information and Callable Credit Default Swaps" visualizes a specific opportunity and challenge of the introduction of a credit derivatives market. Credit derivatives provide banks with the opportunity to transfer credit risk beyond lending constraints to external investors. The challenge, however, is that standard credit risk transfer also introduces information asymmetry costs. If a bank can not credibly signal loan quality, the transfer negatively affects its incentives to screen and monitor the underlying loans which causes overinvestment, or a market breakdown. I show that simple structuring of credit derivatives solves this problem. The basic idea is that credit risk should be transferred by using a callable credit default swap. As the bank can signal loan quality by expressing its readiness to pay for the implicit call feature, screening and monitoring incentives are maintained. The main insight of this paper is that while credit derivatives can have a negative impact on the economy by undermining screening and monitoring incentives, adequately structured credit derivatives allow to solve these asymmetric information problems. The paper also illustrates that stricter regulatory capital requirements may impede effective signaling mechanisms. This argument is often neglected in the current discussion on how regulatory capital could help to support financial stability.

The second paper with the title "Macroeconomic Conditions, Growth Options and the Cross-Section of Credit Risk" is joint work with Alexander Wagner and Ramona Westermann. It considers the pricing of debt if both firm-specific risk and macroeconomic risk are incorporated. The central new feature of our paper compared to the literature in this field, which only considers invested assets, is that firms are composed of both invested assets and growth options. We show that firms with a high portion of growth options in the value of their assets have larger costs of debt for two reasons. First, options are more volatile than assets in place. Second, because they lose more value than assets in place when the economy switches to recession, options induce a higher tendency to default during bad times when marginal utility is high and recovery rates are low. Exploring this insight allows us to explain stylized facts regarding empirically observed credit spreads, default probabilities, and leverages. In particular, we show that growth options explain the credit spreads puzzle, i.e., the fact that standard credit risk models calibrated to historical parameters typically underestimate the credit spreads observed in reality. Because our model features firms which are different with respect to the importance of growth options in the value of their assets, we can also explain the empirical cross-section of both credit spreads and leverage. Additionally, the results are consistent with observed pro-cyclical aggregate investment spikes and busts, and with counter-cyclical default clustering. Overall, we suggest that it is crucial to consider the impact of growth options when modeling corporate credit risk and the impact of this risk on firms.

The effect of cash holdings on credit risk is analyzed in the last article "The Impact of Managerial Control over Cash on Credit Risk and Financial Policy". By introducing a cash policy and manager-shareholder conflicts into a trade-off model of capital structure, I identify two channels through which cash affects the credit risk of firms. First, the fact that managers can use cash to service debt when equityholders are unwilling to inject funds into the firm reduces credit risk. Second, as equityholders anticipate that ceasing to inject funds does not lead to immediate firm default, they optimally stop contributing funds earlier in firms with cash holdings compared to a firm without cash. The second channel increases credit risk. Because the two channels depend on individual firm and industry characteristics, the insights induce - much like in the previous paper - conclusions not only about the credit risk of a typical firm, but also about the cross-section of credit spreads. Additionally, as managers target excess cash to reduce the probability of firm default, the insights on the effect of cash on credit risk allow to discuss a wide range of stylized facts regarding corporate cash policy choices. In particular, I explain why managers of firms with larger distress costs, a higher yield of liquid assets, higher asset volatility, or higher takeover costs target a larger amount of excess cash. Hence, the results of the paper shed light on both credit spreads of firms with cash holdings and potential agency conflicts between managers and equityholders regarding the corporate cash policy.

References

Batten, Jonathan A., and Warren P. Hogan, 2002, A perspective on credit derivatives, *International Review of Financial Analysis* 11, 251–278.

Chen, Hui, 2010, Macroeconomic conditions and the puzzles of credit spreads and capital structure, *Journal of Finance* 65, 2171–2212.

Duffee, Gregory R., and Chunsheng Zhou, 2001, Credit derivatives in banking: Useful tools for managing risks?, *Journal of Monetary Economics* 48, 25–54.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, 2001, Explaining the rate spread on corporate bonds, *Journal of Finance* 56, 247–277.

Huang, Jing-Zhi, and Ming Huang, 2003, How much of the corporate-treasury yield spread is due to credit risk? A new calibration approach, *Mimeo*.

Instefjord, Norvald, 2005, Risk and hedging: Do credit derivatives increase bank risk, *Journal of Banking and Finance* 29, 333–345.

ISDA, 2010, International swaps and derivatives association market survey 1987-2010, *Mimeo*.

PART II: RESEARCH PAPERS

Private Information and Callable Credit Default Swaps

August 25, 2011

Abstract

This article analyzes the impact of alternative credit risk transfer mechanisms on bank behavior. A bank can screen and monitor new loans to collect private information. While standard credit derivatives allow to transfer credit risk to investors, they negatively affect banks' incentives to screen and monitor. The reason is that a bank cannot credibly signal private information on loan quality to external investors. Fortunately, simple structuring of full protection credit derivatives solves the information asymmetry problem. In particular, it is shown that a callable credit default swap reveals a loan's quality to the investor by letting him observe the bank's readiness to pay for the call feature. This signal restores incentives for beneficial screening and monitoring. The paper also examines the influence of regulatory capital requirements on incentives, signaling, and the design of the optimal credit risk transfer contract.

1 Introduction

The credit derivatives market segment has been one of the most innovative and fastest-growing before the breakout of the recent subprime mortgage crisis. Barrett and Ewan (2006) estimate a market size of USD 5 trillion for 2004. Two years later, the estimated outstanding notional already amounts to USD 20 trillion, exceeding the US GDP of USD 13 trillion in 2006. Despite the tremendous growth of the credit derivatives market, these instruments have so far hardly been extended to transfer the risk of bank commercial loans. One reason, according to Duffee and Zhou (2001), is that information asymmetry among players in the market presents a major challenge. Academics and practitioners argue that credit risk transfer weakens banks' credit discipline in the presence of information asymmetry, which might be one of the central reason for the subprime mortgage crisis. The dramatic decrease in activity on structured credit derivative markets in 2008, and the enormous impact of the crisis on the economy painfully reveal the importance to solve the problem.

This paper begins by showing that credit risk transfer adversely affects the screening and monitoring incentives of lenders. The theoretical results confirm the empirical findings in Keys et al. (2008), and Ashcraft and Sanots (2009). The disincentives induce higher costs of debt financing, which in turn causes economic costs. Building on these insights, a new and distinct way to overcome the problem is introduced, namely callable credit default swaps. Subsequently, I discuss the optimal design of credit derivative contracts by comparing the proposed solution to partial protection approaches described in the literature. It turns out that the optimal security design depends on regulatory capital requirements. While designed to sustain stability in the financial system, these requirements in fact often impede credible partial protection solutions. As a consequence, callable credit default swaps evolve as the optimal signaling contract in most cases.

The basic situation I study is the following. Consider three market participants, a bank (B), an investor (S), and a borrowing firm (F). Their relationship is sketched in Figure 1. Suppose that the loan-originating B is subject to credit risk constraints, either by an internal limit on concentrated lending or by capital adequacy requirements. Thus, B needs to transfer the credit risk of a new loan to S via credit default swap (CDS).[1] S demands a protection fee to compensate the expected costs of credit risk. The problem of standard credit risk transfer techniques is that they reduce B's incentive to screen and monitor the new loan because S, not B, now bears the consequences of adverse loan quality. This misbehavior is anticipated by S, and, hence, reflected in a larger protection fee. B consequently faces information asymmetry costs when transferring credit risk. Additionally, if the protection fee turns out to be above the loan rate earned, the market breaks down as in Akerlof (1970), and an otherwise profitable loan cannot be granted (underinvestment). A bank, therefore,

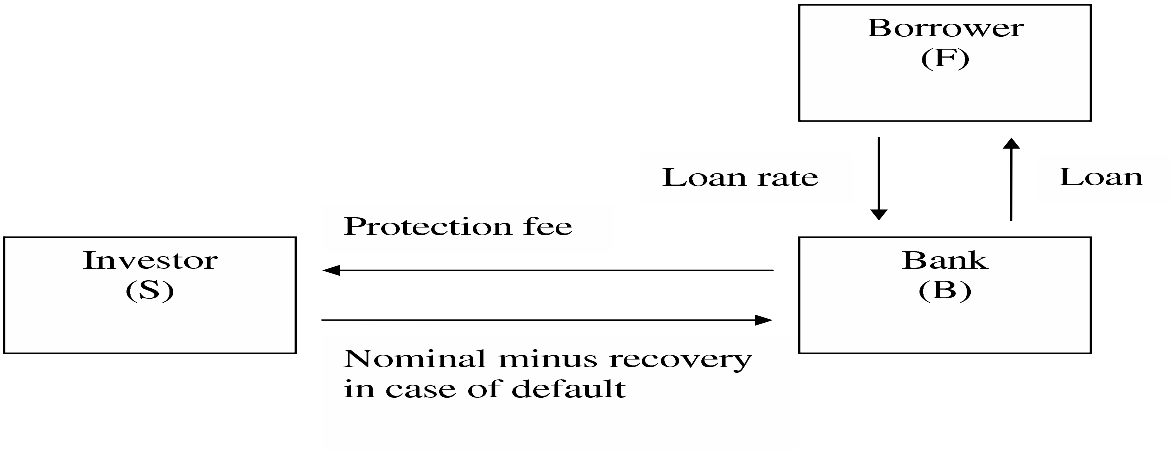

Figure 1: Market participants

faces a basic dilemma: On the one hand, relationship advantages suggest granting loans to well-known borrowers. On the other hand, undiversified lending requires credit risk transfer to reduce a concentration of risk, which induces information asymmetry costs. Relaxing this trade-off by decreasing information asymmetry costs is desirable for several reasons: First, Von Thadden (1998) demonstrates that loans granted to relationship-borrowers increase profitability in credit markets due to accumulated information. Second, as Duffee and Zhou (2001), and Duffie (2007) point out, the positive effect on the economy, and particularly on banks, is likely to be large if credit derivatives penetrate the market of formerly illiquid, bank-originated credit risk, and contribute to a more efficient distribution of credit risk in

[1]A credit default swap is a specific kind of counterparty agreement which allows the transfer of third party credit risk from one party to the other. For example, one party in the contract could be a lender facing credit risk from a third party borrower, and the counterparty in the credit default swap agrees to insure this risk in exchange of regular periodic fee payments.

the economy.

I argue that credit derivatives' flexibility in repackaging risk allows to restore screening and monitoring incentives, thereby lowering the fair protection fee and preventing a market breakdown. Following Duffee and Zhou (2001), B's informational advantage is assumed to be relatively small for short-term payoffs, but relatively large concerning the payoffs far in the future. In this situation, consider using a callable CDS (CCDS) contract to transfer the credit risk to S. A CCDS is a credit default swap which can be canceled at a predetermined date by the protection buyer. Once this date is reached, and B decides whether to call the contract, the bank does not have an informational advantage any more because only short-term payoffs remain up to maturity. Hence, the informed B's readiness to pay for the implicit call feature constitutes a credible signal of the loan's quality ex ante. The signal, therefore, allows to express screening and monitoring effort. B has an incentive to engage in these activities to lower the protection fee. It thereby reduces information asymmetry costs, while still achieving the primary objectives of the credit derivative transaction, namely diversification, optimization of economic and regulatory capital, and complete risk transfer. This logic suggests that by fine-tuning simple credit derivatives, banks can ultimately solve the basic dilemma, i.e., extend loans to well-known borrowers while transferring excessive credit risk without information asymmetry costs.

The article relates to a variety of strands of the literature. First, it is based on the broadly discussed idea of banks having a unique ability to build relationships with their borrowers, thus simplifying monitoring (Diamond (1984)), long-term commitment (Von Thadden (1995)), and screening. There is a substantial debate among academics and practitioners about the effects of credit derivatives on bank behavior and the bank-borrower relationship. Concerns about credit derivatives undermining positive relationship-rents by causing misbehavior are expressed in Kiff and Morrow (2000) and Morrison (2005). For example, Morrison (2005) argues that credit derivatives could adversely affect banks by reducing their incentive to screen and monitor borrowers. The use of credit derivatives may render bank loans less valuable because the loans would entail less of a certification effect. The current article adds to this debate and states that properly structured credit derivatives do not erode the rents generated in the bank-borrower relationship. Second, the paper relates to the discussion among Von Thadden (1995), Gale and Hellwig (1985), Innes (1990), and others on the application of strategic contracting within financial intermediation to mitigate information asymmetry problems.

For credit markets, the lemons problem and the ability to sell loans if banks have private information are discussed in Carlstrom and Samolyk (1995). In their setting, buyers realize that banks are selling loans due to capital constraints. Hence, the former acquire exposures even when they cannot perfectly screen the ex ante quality of loans, whereby the standard lemons problem can be avoided. The classical reference to loan sales and information asymmetry is the paper of Gorton and Pennacchi (1995). They conclude that if a bank can implicitly commit to holding a certain fraction of a loan, i.e., to provide limited recourse, the moral hazard associated with the loan sales market is reduced. Similar ideas are subsequently applied to articles on credit derivatives structured to mitigate information asymmetry problems. The first paper which rigorously considers the implications of credit derivatives for banks' risk-sharing is Duffee and Zhou (2001). The authors show how banks hedging high-quality loans can use credit derivatives with a maturity mismatch[2] to shift the risk of early default to outsiders. By retaining the risk of late default they avoid the lemons problem. Boot et al. (1993) provide the basic idea of splitting a risky cash flow into a senior and a subordinated security: The senior security is considered to be information-insensitive and can be sold to uninformed investors while the subordinated security is information-sensitive and, hence, tailored to informed investors. Riddiough (1997) extends this reasoning by arguing that loan bundling admits pool diversification which softens information asymmetry. DeMarzo and Duffie (1999) show that pooling and shearing of loans allows the protection buyer to concentrate the "lemon's premium" in the first-loss block. Retention of this informationsensitive block reduces the total lemon's premium by aligning the interests of the protection buyer with those of the protection seller. A variety of papers follows the same idea. Franke et al. (2007), for example, model the optimal design of Collateralized Debt Obligations (CDOs). According to Nicolo and Pelizzon (2008), first-to-default credit derivatives and binary credit

[2]In a maturity mismatch, the maturity of the credit derivative contract does not match the underlying loan contract.

default basket contracts can be designed in a similar way to signal the ability of banks to screen their borrowers.

The approaches discussed in the literature concentrate on signaling the bank's type by varying the quantity of insurance. This solution is a standard result within the insurance theory.[3] However, credit risk transfer can not entirely relax a bank's lending constraint whenever signaling requires to retain part of a new loan's risk. In contrast, this paper shows how to signal a loan's type even though the underlying credit risk is completely transferred to an investor.

It is important that the credit risk can be transferred completely in the signaling game. The reason is that market opacity prevents banks from credibly committing to retain part of a loan's risk: A bank can transfer the remaining risk silently without informing either party. In fact, the current regulatory treatment in the Basel II jurisdiction may even encourage banks to do so in order to avoid the regulatory costs incurred with partially retained credit risk. An investor, consequently, does not know whether a bank truly retains some risk exposure and the corresponding incentives to screen and monitor a loan.[4] I argue that, in contrast to partial retention contracts, CCDSs provide a credible signal even if credit derivative trades are private in line with current market practice. As CCDS fully transfer the credit risk of a loan, there are no regulatory retention costs and, hence, no regulatory incentives to silently sell the implicit call feature.

The structure and the price of the credit derivative are the only information required to signal quality to the investor. The signal does not rely on generally unavailable information such as whether the lending constraint is binding, the reason why a bank is selling credit risk (Carlstrom and Samolyk (1995), Nicolo and Pelizzon (2008)), or hard-to-judge reputation effects (Gorton and Pennacchi (1995)). Finally, the choice of a premium as a signal yields an additional striking feature: Standard signaling models often use a wasteful signal.[5] In

[3]"Good" banks signal their quality by buying less insurance. "Bad" banks prefer to buy full insurance and to reveal their type.

[4]In line with this argument, there was a very active market for first-to-default tranches before the current crises, which allowed banks to easily sell retained first loss pieces.

[5]In Spence (1973), education is a wasteful signal for the participants in the model. Partial coverage in Franke and Krahnen (2005), and Duffee and Zhou (2001) is wasteful if costs arise due to underdiversification.

contrast, the signaling premium provided in this paper accrues to the investor, an argument enhancing the marketability and liquidity of credit risk.

For tractability, my model simply addresses two types of loans and two periods. Increasing the number of types to a continuum, or varying the length of the two subperiods does not change the basic insights of this paper as long as the structure of the asymmetric information is maintained. Furthermore, the proposed signaling-mechanism is more beneficial the more severe the information asymmetry problem.

The paper is organized as follows: The next section introduces the structure of the model. Section 3 discusses adverse selection, i.e., the screening problem, and derives the optimal contract. In an extension in Section 4, I show that CCDSs increase the monitoring effort of an intermediating bank, before I combine adverse selection and moral hazard. Section 5 discusses the results and concludes. All valuation techniques and proofs are relegated to the Appendix.

2 A simple model of adverse selection

2.1 Structure of the model

Consider a risk-neutral, profit-maximizing bank B providing a borrower F with a loan. B is operating in a specific loan market and maintains a close relationship to F. The bank's incremental costs to screen and monitor the borrower are, therefore, relatively low. If internal or external constraints on concentrated credit risk are binding, the bank seeks to completely transfer the credit risk of additional loans to an investor S. Financial markets are assumed to be competitive, and the risk-free interest rate is zero.

B's problem is to maximize expected profits out of a given loan and the corresponding credit derivative. To transfer the credit risk of the borrower's loan to S, B chooses between a credit default swap (CDS) and a callable credit default swap (CCDS). The structures do not differ with respect to the convention that S covers the loss of the reference credit's face value following a credit event[6] in exchange for a fair protection fee. A (European) CCDS, however, gives B the additional right but not the obligation to call (unwind) the contract at a predefined point in the future. The structure can be split into its generic components, namely a CDS and an embedded receiver default swaption (RS)[7] with corresponding strike and exercise date. There are three dates in the model outlined in Figure 2.

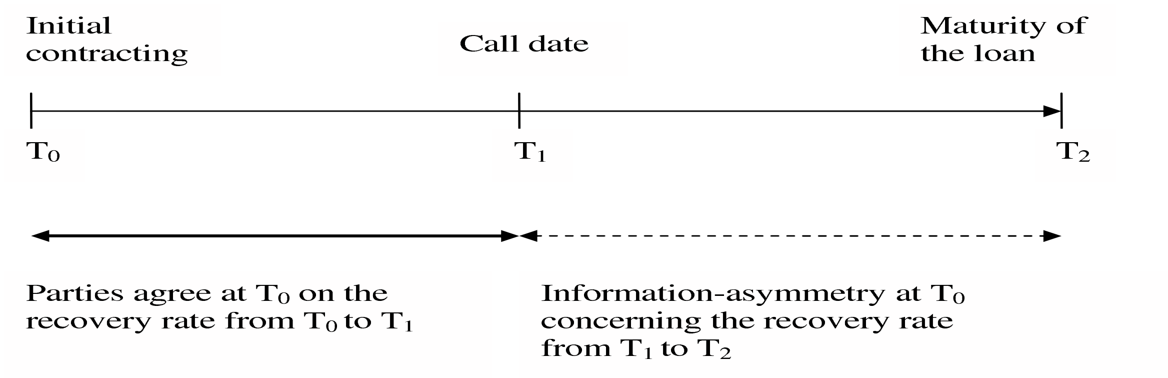

Figure 2: Time structure

T_0 is the starting date, T_1 the call date of the CCDS, and T_2 the maturity date of the loan. At T_0, B decides whether to screen a loan applicant based on public information. Screening at costs C allows to gather private information which reveals the true CDS rate of F's loan. On the basis of this information, B then offers a CDS or a CCDS to S. Without screening, this offer is based on public information. As soon as S signs the credit derivative contract, the bank can grant the loan because the lending constrained is relaxed. I only consider one call date at T_1. In a CDS, the protection fee payable by B to S is fixed during the contract's lifetime from T_0 to T_2. Transferring the credit risk of a loan via CCDS includes an option for B to call the protection at time T_1. I assume that given its credit risk constraint, the bank needs to hedge the loan with a CDS after a call. It is, consequently, only worthwhile for B to call if it can buy new protection at T_1 up to T_2 for a lower fee than the one initially agreed.

The information structure of the model deserves a closer description. I distinguish between freely available public information and costly private information. The former contains loan

[6]A credit event is a legally defined event which typically includes bankruptcy, failure to pay and restructuring. Note that the correlation between a credit event of B and S is assumed to be zero.

[7]In a receiver default swaption, the option buyer pays a premium to the option seller for the right, but not the obligation, to sell CDS protection on a reference entity at a predetermined rate (strike) on a future date. See O'Kane et al. (2003), page 26.

pricing parameters including the recovery rate R[8] for defaults during the *next* time interval. To determine the fair CDS rate at T_0 for protection up to maturity, however, one does not only need the recovery rate from T_0 to T_1, but also the one from T_1 to T_2. The latter piece of information corresponds to the private information. Loans with a high R from T_1 to T_2 (high loans) yield a fair CDS rate s^H, and loans with a low recovery rate (low loans) have a fair CDS rate s^L. These fair CDS rates compensate an investor for the credit risk of a loan. Consistent with market practice, the investor is unable to observe a bank's credit derivative contract offers to other third parties, the bank's lending constraints, and the screening and monitoring activities.

Figure 2 shows how the structure of the asymmetric information varies over the life of the loan. The information asymmetry at T_0 between S and a bank with private information refers to the recovery rate over the future time interval from T_1 to T_2. The key observation to approach adverse selection is the following: B and S realize at T_0 that, once they reach time T_1, there will be no more information asymmetry, since the recovery rate for the proximate period is public information.

Lacking private information, S observes B's choice of the hedging strategy, and eventually updates beliefs concerning the recovery rate of a loan. Let μ denote the probability of a high loan if beliefs are based on public information. ρ indicates the assessed probability of a high loan if beliefs are updated after observing the bank's contract offer. The spreads $s(\mu)$ and $s(\rho)$, respectively, then denote the *expected* CDS rates. They represent the fair protection fee demanded from S for bearing the loan's credit risk.

A fixed, risky loan rate i is charged to the borrower F for the loan. Administrative or operating expenses of the bank can simply be incorporated by reducing i. The bank's payoff from using CDSs is formulated as follows:

$$Max\left[0, V_{0,2}^{fee}\left(i - s(\mu)\right)\right] \tag{1}$$

$V_{0,2}^{fee}$ is the present value at T_0 of receiving one basis point of fee payments up to T_2 as

[8] The recovery rate is the value of a loan at default.

long as there is no default. Details on pricing can be found in the Appendix. B grants and hedges a loan if the overall expected profit is positive ($V_{0,2}^{fee}\big(i - s(\mu)\big) \geq 0$). Rejection results in a present value of zero. The protection fee $s(\mu)$ payable to S is based on public information of μ. It is fixed at time T_0 up to T_2.[9]

In contrast to CDSs, CCDSs allow to exploit the time varying structure of the information asymmetry. Consider the optimization problem of a bank that has detected a high recovery credit. The profit of the bank corresponds to

$$\underset{P}{max}\big[V_{0,2}^{fee}(i - s(\rho)) - P + A^H - C\big]. \tag{2}$$

B earns the rate i from the loan and pays a protection fee $s(\rho)$ to S. P denotes the premium paid for the right to call the contract. The economic value of this right is known to B after screening and corresponds to A^H for a high, and A^L for a low loan. It depends (besides publicly known parameters) on the recovery rate R from T_1 to T_2, i.e., on the private information. B and S know at T_0 that the contract is eventually called at T_1. As there will be no more information asymmetry at this date, the value of the call feature to B does not depend on the private information. Hence, a bank is able to signal the loan's type by expressing its readiness to pay for the receiver swaption at T_0. In particular, after screening and detection of a high loan, it simply offers a larger P than the premium a bank with a low recovery credit - or a bank which has not screened - is ready to pay. As a consequence, only high loans are hedged with CCDSs. S, in turn, updates beliefs to $\rho = 1$ which induces $s(\rho) = s^H$. The readiness to pay for the implicit call feature at time T_0 signals the future recovery rate, and dissolves information asymmetry.

In the signaling game between B and S, the latter is not perfectly informed about the loan. Therefore, I follow Osborne (2004) and use the model of a strategic, Bayesian two-player game with imperfect information. A pure strategy equilibrium is defined as a duple of actions, one for the investor S and one for the bank B. The actions of each player are

[9] It seems like repeatedly signing short-term credit default swaps for proximate periods is a solution to the problem. Unfortunately, this procedure does not relax the credit risk constraint, as the underlying loan is not fully hedged in a maturity mismatch.

the following: B can choose to grant and hedge a loan without screening, to stay out of the market if the expected profit is negative, or to screen and select the appropriate instrument to transfer the risk of the exposure, i.e., a CDS or a CCDS. In an extension, the bank may also monitor the loan after the risk has been transferred. S can either reject or accept the credit derivative offer. The payoff to each player depends on the other player's action and on the market environment. In an equilibrium of the Bayesian game, the action chosen by each player is optimal, given the action chosen by the other player. An equilibrium is explicitly defined with respect to the perceived probabilities of high and low loans.

2.2 Discussion of the assumptions

The model assumes that S cannot infer the credit's type by looking at the loan rate. This assumption is not unrealistic, given that (i) B may not have an incentive to screen the loan to determine the fair loan rate, (ii) the asymmetry of information affects the loan rate charged by the bank[10], (iii) banks are competing for borrowers[11], (iv) the credit is merely one part of the overall relationship between the bank and the borrower, (v) the bargaining power of counterparties and the market structure play a role in the determination of the loan rate as argued in Petersen and Rayan (1995). Rather than explicitly modeling these aspects, I simply assume that the borrower F pays a fixed loan rate i, and, hence, that i is an imperfect signal of loan quality.[12]

Assuming that screening gives a market participant an information advantage for long maturities, but not for short maturities, deserves some motivation. The intuition is linked to the one given in Duffee and Zhou (2001). Investments of a borrowing firm and their performance are publicly observable during the life of the loan, but *planned* investments are

[10]Stiglitz and Weiss (1981) agree on the opaque relation between borrower quality and loan rate. They assume an asymmetric information problem between the borrower and the bank. The form and magnitude of this asymmetry affects the loan rate charged.

[11]This argument is provided by Von Thadden (1998): In order to attract borrowers, low loan rates are offered, leading to expected losses in the short term. Over time, by building up a relationship, information about the borrower is accumulated, creating an information advantage over other lenders. This information allows a relationship bank to extract informational rents from the borrower at a later stage, since the former is able to issue tailored counteroffers for its most valuable customers based on its inside information.

[12]Note that I do not incorporate information asymmetry between the borrower and the bank. If the borrower knows his own loan type, screening costs are redundant given that debt contracts exist which reveal the borrower's type to the bank without relying on a costly signal.

not. If default occurs early after contract implementation, the recovery rate mainly depends on existing assets as the firm has hardly invested in new, planned projects. The parties are, therefore, likely to agree on the recovery rate of early defaults. In contrast, the recovery rate of a default far in the future heavily depends on future planned projects. As long as screening releases information about such plans, the resulting informational advantage mainly refers to long maturities.

F's *planned* projects are indeed difficult to assess for market participants. Due to the close relationship to the borrower[13], guarantee of confidentiality[14], offered consulting and expertise, and F's access to various bank services, B is in a privileged position to acquire such information. Evaluating or influencing the recovery rate over a future interval further requires that planned projects, netting agreements, collateral arrangements, and securities held by the bank are analyzed. Therefore, it is reasonable to assume relatively low screening and monitoring costs for B compared to outside investors.

In light of the results of Gupton et al. (1997), it makes sense to consider the future recovery rate as the private information. They demonstrate that there is wide variation in recovery rates even for the same subordination level, which induces that the recovery rate is one of the most uncertain input parameter in a pricing model. Moreover, Schönbucher (1999a) shows why different perceptions of recovery rates have a relatively high impact on prices of CDSs. Especially for high default probabilities, disparities in recovery rates result in profound differences of CDS spreads.

3 Analysis

3.1 Overview of the results

The outcome of the analysis of adverse selection depends on the level of the loan rate. Figure 3 characterizes four possible scenarios.

[13]A formal or informal relationship helps to impose pressure upon the borrower or to perform pareto-improving renegotiations (Gorton and Kahn (2000)). In addition, a relationship can form an implicit contract

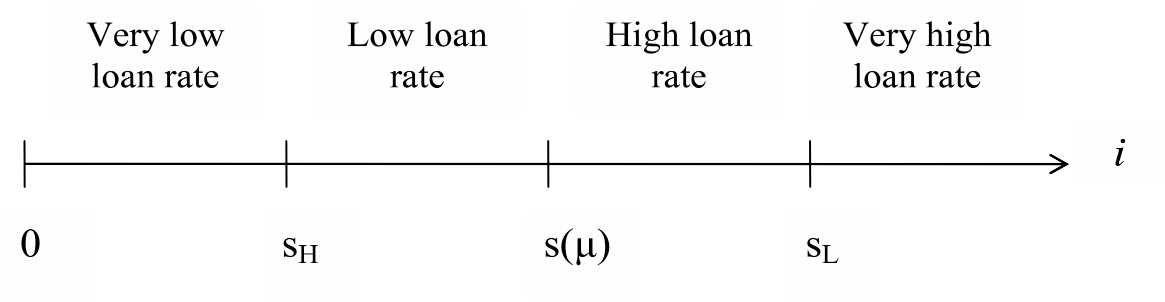

Figure 3: s_L denotes the fair CDS rate of a loan with a low recovery rate, s_H the one of a loan with a high recovery rate, and $s(\mu)$ is the expected CDS rate based on beliefs μ expressing the public information on the probability of a high loan. The figure shows four possible scenarios for the loan rate i: A very low loan rate is characterized by $i< s_H$, a low loan rate by $s_H \leq i< s(\mu)$, a high loan rate by $s(\mu) \leq i< s_L$, and a very high loan rate by $i\geq s_L$.

In what follows, I illustrate the low loan rate scenario in detail. Table 1 provides an overview of the results. Results for the remaining scenarios are discussed in Section 3.5.

	Condition for screening	**Expected bank profit**	**Overall welfare**
First Best	$\mu(i-s^H) \geq C/V_{0,2}^{fee}$	$\mu(i-s^H)V_{0,2}^{fee} - C$	$\mu(i-s^H)V_{0,2}^{fee} - C$
CDS	No screening	0	0
CCDS ($SP > 0$)	$\mu(A^H - A^L) \geq C/(1-\mu)$	$\mu(i-s^H)V_{0,2}^{fee} - C$ -SP	as in the First Best
CCDS ($SP = 0$)	as in the First Best	as in the First Best	as in the First Best

Table 1: A^H and A^L denote the fair value of the call feature on a high, and a low recovery loan, respectively. C expresses screening costs, $V_{0,2}^{fee}$ is the present value of receiving one basis point of fee payments up to T_2 as long as there is no default, and SP is the signaling premium. The table summarizes the results of credit risk transfer for low loan rates in four cases: (i) without information asymmetry (First Best), (ii) with information asymmetry using a CDS (CDS), (iii) with information asymmetry using a CCDS given a positive signaling premium (CCDS($SP > 0$)), (iv) with information asymmetry using a CCDS without a signaling premium (CCDS($SP = 0$)). For each case, the table outlines the condition for screening taking place (condition for screening), the expected bank profit given that the condition for screening is satisfied (expected bank profit), and the combined expected profit of the bank and the investor given the condition for screening is satisfied (overall welfare).

3.2 The First Best in an environment of low loan rates

The analysis in this section assumes a *low loan rate*, $s_H \leq i< s(\mu)$, i.e., that the loan rate is higher or equal to the CDS rate of a high recovery loan, but smaller than the expected CDS

regarding borrowing and repayment beyond the formal explicit legal contract on which it is based.

[14]Bhattacharya and Chiesa (1995) show how confidentiality of a bank may encourage its clients to reveal more information.

rate under μ. The First Best is characterized by a market without information asymmetry among the buyer and the seller of credit risk. It represents a situation where a bank is able to bear an additional credit risk of a loan itself, or an environment without informational frictions on the credit risk transfer market. A bank is able to assess the recovery rate of a new credit for defaults up to maturity. Once this costly information is produced, it becomes common knowledge. Without information asymmetry, a high or low recovery loan can then be hedged for a protection fee corresponding to the fair CDS rate s^H or s^L, respectively.

In what follows the term "Net Present Value of a loan" (NPV) is used:

Definition 3.1. *The NPV of a loan from T_0 to T_2 is given by $V_{0,2}^{fee}(i-s^H)$ for a high recovery credit, and by $V_{0,2}^{fee}(i-s^L)$ for a low recovery credit. It is positive if the loan rate is higher than the fair CDS rate (good loan), and negative otherwise (bad loan). Whenever the credit risk is transferred, the NPV of the loan to B is given by $V_{0,2}^{fee}(i-s(\cdot))$, where $s(\cdot)$ represents the protection fee.*

A bank exclusively relying on public information always turns down a new loan in the low loan rate scenario, resulting in a market breakdown with zero expected profit. The reason is that $s(\mu)$ is higher than the loan rate, yielding a negative NPV for granted loans. However, a market breakdown induces underinvestment, since good loans are also rejected. Instead of accepting this outcome, it may be worthwhile for a specialized bank to screen F at costs C. Once B knows whether being faced with a good or bad loan, it grants the former and rejects the latter. B prefers to screen at costs C if the expected profit of doing so, i.e., $\mu(i-s^H)V_{0,2}^{fee}-C$, is greater or equal than in a market breakdown.

Proposition 1. *For a low loan rate, the bank screens a new loan applicant in the First Best as long as $\mu(i-s^H)\geq C/V_{0,2}^{fee}$.*

Proof. See the Appendix. □

The analysis reveals that the ability to screen loans is valuable as long as screening costs are low because it generates positive expected profits in an otherwise unprofitable market.

The first row of Table 1 shows the above derived condition for screening, the expected bank profit, and overall welfare. Overall welfare incorporates the total expected profit of the loan and the CDS contract to B and S.[15] In a competitive market without information asymmetry, the CDS is a zero NPV contract. Hence, the overall welfare from the loan corresponds to the expected bank profit.

3.3 CDSs as a hedging tool in an environment of low loan rates

If a bank keeps on lending to well known borrowers it will reach a point where internal or external lending constraints are binding. Constraints can be relaxed by credit risk transfer techniques.

When S decides whether to accept the terms of a CDS, he does so on the basis of some probability assessment over the set of types of loans underlying the credit derivative contract. The minimum protection fee for which a risk neutral investor S with public information is willing to sign a contract hedging the counterparty against the credit risk of a loan is

$$s(\mu) = \mu s^H + (1-\mu)s^L. \tag{3}$$

If B argues that the underlying loan is of high recovery type, S may update beliefs to ρ, resulting in a protection fee of $s(\rho) = \rho s^H + (1-\rho)s^L \leq s(\mu)$. A bank will, however, always claim to be hedging a high recovery credit in an attempt to decrease the protection fee, even if it has detected a low recovery credit. There are no costs of doing so. Without a possibility for B to credibly signal the loan's type, costly screening is useless because it does not allow to decrease the protection fee by influencing beliefs. As S anticipates that screening does not take place in an equilibrium, he continues to believe that a loan is of high type with probability μ and of low type with probability $(1-\mu)$. Given the bank's strategy, S accepts a contract with a CDS rate equal to $s(\mu)$ to cover the expected credit risk. Hence, no new loans are granted as $i < s(\mu)$. The Bayesian equilibrium with CDSs is a market breakdown without screening.

[15]The model in this paper does not account for additional benefits to F out of a granted loan.

Proposition 2. *Information asymmetry between the bank B and the investor S impedes screening activities, if lending constraints are relaxed by transferring the credit risk via CDSs. The loan market breaks down, resulting in underinvestment.*

Proof. See the Appendix. □

Conceptually, the result for the low loan rate scenario is in line with Gale (1990) and Hubbard (1998): With asymmetric information, the credit market does not fund all socially efficient projects. The underinvestment problem is evident in Table 1: Recall that in the First Best, a loan is screened as long as the screening-condition is satisfied, yielding a positive overall welfare and expected bank profit. Credit risk transfer exacerbates the situation, resulting in zero profits for all participants in the market.

The advantage of a CDS-market, i.e., the ability to grant profitable loans even though credit risk constraints are binding, is diluted by screening disincentives due to information asymmetry. This friction causes costs, as profitable credit markets break down. Banks, therefore, clearly have an incentive to mitigate the investors' informational disadvantage.

3.4 CCDSs as a hedging tool in an environment of low loan rates

CDSs provide no signal of loan quality which causes a market breakdown in case of low loan rates. This section explains how structuring the hedge in a simple way allows to mitigate information asymmetry costs, and, hence, to prevent a market breakdown.

Let A^H express the fair value of the call feature on a high recovery loan, and A^L the one on a low recovery credit. They correspond to the value of a receiver swaption (RS) on the CDS rate of a high, or low loan, respectively, with exercise date T_1, maturity T_2, and strike s^H.[16] The pricing formula for the RS is given in the Appendix. The following lemma describes the impact of the recovery rate on the value of a receiver swaption:

[16] B can buy protection via CDS and synthetically create the possibility to call this contract (replicate a CCDS) by entering a RS. If the RS is exercised at T_1, then B additionally *sells* protection up to T_2. Combined, the CDS and the exercised RS exactly offset each other. A CDS is synthetically called (unwinded) at T_1 in this way. One is, therefore, able to separate the value of a call feature as the value of a RS.

Lemma 3.2. *The value of a receiver swaption is increasing in the recovery rate:*

$$\frac{\partial RS}{\partial R} > 0. \tag{4}$$

Proof. See the Appendix. $\square$

Lemma 3.2 allows B to signal the underlying loan's type by expressing the readiness to pay for the implicit call feature of a CCDS. For an equilibrium, one first needs to ensure that S participates in the credit risk transfer contract. His participation, or individual rationality constraint (IR) requires that

$$P \geq A(\rho), \tag{5}$$

where P denotes the price paid for the call feature, and $A(\rho)$ is the value of this feature to S under beliefs ρ.

Next, I discuss how beliefs are formed. After screening and detection of a high recovery loan, B is offering a certain call-premium P to S. The latter is simply chosen such that (i) a bank which has detected a low recovery credit weakly prefers to stay out of the market, and (ii) a bank weakly prefers to screen and subsequently make an appropriate contract choice than to pick the CCDS contract without screening at all.[17] B then selects a CCDS contract if a high loan has been detected, or to stay out of the market in case a low loan has been identified. (i) is satisfied if $iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - A^L) \leq 0$, while (ii) requires that $iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - \mu A^H - (1-\mu)A^L) \leq \mu\big(iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - A^H) - C$. As (ii) is more restricting than (i), the following incentive compatibility constraint (IC), obtained by simplifying (ii), expresses both conditions:

$$(i - s(\rho))V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq P. \tag{6}$$

To minimize credit risk transfer costs, a bank with a high loan is choosing the smallest call-

[17] (ii) is necessary because S cannot observe the screening-activity itself.

premium P* simultaneously satisfying the IR and the IC. If P* still admits a positive expected profit for B, S updates beliefs to ρ as follows: "The underlying loan is of high recovery type with probability one ($\rho = 1$, and $s(\rho) = s^H$), since B would not offer P* if it had not screened and detected a high recovery credit."

Third, even if B is endowed with a mechanism to credibly express loan quality, it may not be worthwhile to participate in the market. The last step, therefore, is to confirm whether the CCDS approach increases expected profits compared to the outcome with CDSs:

$$\mu\big((i - s(\rho))V_{0,2}^{fee} - P^* + A^H\big) - C. \geq 0 \tag{7}$$

$(i - s(\cdot))V_{0,2}^{fee}$ corresponds to today's value of the loan rate payments minus the protection fee payments up to default or maturity. The left hand side of Inequality (7) shows the expected profit of a bank choosing to screen at costs C. With probability μ, a high recovery credit is detected. B then signals quality by paying P* in exchange for a call right value A^H. With probability $(1-\mu)$, screening reveals a low recovery credit. B prefers to stay out of the market in this case, because misleadingly signaling a high recovery credit by paying P* would generate an expected loss (see the IC). The right hand side of Inequality (7) represents the market breakdown, which occurs if only CDSs are available. Plugging the lowest P* satisfying the IR and the IC into Inequality (7) leads to the result in the following proposition:

Proposition 3. *The intermediating bank prefers to screen and use CCDSs to the outcome with CDSs, if and only if*

$$\mu(A^H - A^L) - \frac{C}{(1-\mu)} \geq 0. \tag{8}$$

Whenever $(i - s^H)V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H$*, the Screening Condition (8) reduces to* $\mu(i - s^H) \geq C/V_{0,2}^{fee}$ *as in the First Best.*

Proof. See the Appendix. □

In contrast to standard credit risk transfer techniques such as CDSs, CCDSs allow to signal loan quality. As a consequence of the possibility to lower the protection fee, B is induced to collect private information on loans if C is not too high. Therefore, underinvestment can be avoided even though the credit risk is fully transferred to S: By screening and subsequent intermediation of good loans, profitable market activity is maintained. The Appendix shows that this result is not affected by market opacity: Even if credit derivative contracts and the bank's lending constraint are not publicly observable, the basic mechanism maintains. The reason is that the price paid at contract initiation contains the signaling costs. Hence, they are irreversible and can not be silently avoided after the initial contract date.

For a comparison to the First Best, I define the signaling premium (SP) as the difference between the premium P* and the fair value A^H of the call feature. B's expected profit when using a CCDS depends on this signaling premium. The question is how much above A^H the premium P^* needs to be in order to induce S to update beliefs to ρ, i.e., to satisfy the IC. In the First Best, B maximizes the trade-off between costs of underinvestment and screening costs. Using CCDSs, it optimizes the trade-off with respect to underinvestment, the signaling premium, and screening costs. If $(i - s^H)V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H$, $P^* = A^H$ is sufficient to update S's beliefs. The signaling premium is equal to zero in this case. Facing exactly the same trade-off, B also realizes the same outcome as in the First Best. Hence, CCDSs allow to reach the First Best under certain conditions, as outlined in Table 1 ($SP = 0$). The expected profit, however, is smaller than in the First Best whenever a certain signaling premium needs to be paid, and a more restrictive condition for profitable screening emerges (see $SP > 0$ in Table 1).

3.5 Results for very low, high, and very high loan rates

I first explain the case where the loan rate is very low, i.e., $i < s_H$. As it is publicly known that the loan rate is too low to admit any profitable market activity, no loan is granted or hedged.

Next, consider the very high loan rate case, i.e., $i \geq s_L$, in the First Best environment. Before screening, B relies on public information. It presumes a high credit being detected with probability μ and a low credit with probability $(1 - \mu)$. Screening before granting and hedging a loan is then expected to yield $V_{0,2}^{fee}\big(i - s(\mu)\big) - C$ because both types of credits are granted anyway. In contrast, the expected income increases to $V_{0,2}^{fee}\big(i - s(\mu)\big)$ if B lends without screening. A credit is, consequently, granted without screening in the First Best. The same outcome develops with information asymmetry and CDSs: B has no mean to credibly signal loan quality in a CDS hedge. As a consequence, it omits costly screening. S, anticipating this behavior, assigns beliefs μ to determine the protection fee on offered loans. B's expected income, therefore, amounts to $V_{0,2}^{fee}\big(i - (\mu s^H + (1-\mu)s^L)\big) = V_{0,2}^{fee}\big(i - s(\mu)\big)$ as in the First Best. In the very high loan rate scenario, each loan has a positive NPV and should be accepted. There is no need to spend C, uncover the exact CDS rate, and think about acceptance again. Creating an instrument to induce screening such as CCDSs lacks a benefit, because, keeping overall profitability constant, contractual innovations just cause a redistribution of wealth between the participants. Hence, no screening takes place in the pooling equilibrium, all loans are approved, and the corresponding credit risk is transferred using CDSs.

Finally, the high loan rate case deserves a closer description. A bank generates a positive NPV by just granting credits without screening, because the loan rate is higher than the expected CDS rate. Cross-subsidization[18] of bad credits by good credits allows a positive expected profit. However, screening may still be worthwhile for B if C is low enough. The reason is that spending C, and thereby learning the true recovery rate, allows the bank to reject bad loans, while good ones can still be granted. In other words, cross-subsidization and overinvestment are prevented. Following the proof of Proposition 1, it is easy to show that the bank screens a new loan applicant in the First Best as long as $(s^L - i)(1-\mu) \geq C/V_{0.2}^{fee}$.

Now, consider the credit risk being transferred via CDS in an environment of high loan rates and information asymmetry. S demands a protection fee equal to $s(\mu)$ due to the lack of a credible quality signal. Since screening does not provide a benefit to B, it just grants any loan. The resulting overinvestment causes information asymmetry costs given as the

[18] The term "cross-subsidization" describes the following: For low recovery credits the loan rate is too low, for high recovery credits the loan rate is too high. If the expected profit is positive - and higher than in any screening equilibrium -, the good credits allow the bad ones to be granted without detection. This is possible although the bad credits decrease the total expected profit. It is just more expensive to screen and reject bad loans than leaving them in the market.

difference of expected profits to the First Best. In contrast, CCDSs decrease information asymmetry costs by providing the incentives to screen loans. If C is not too large, the overinvestment problem can be solved even though credit risk is fully transferred to S. The derivation of the corresponding Proposition 4 is analogous to the one of Proposition 3. The only difference is that Inequality (7) needs to be replaced by

$$\mu\big((i - s(\rho))V_{0,2}^{fee} - P^* + A^H\big) - C \geq (i - s(\mu))V_{0,2}^{fee}, \tag{9}$$

because the bank's expected profit with CDSs corresponds to $(i - s(\mu))V_{0.2}^{fee}$ in case of high loan rates.

Proposition 4. *Consider a market characterized by a high loan rate. A pooling equilibrium with loans being hedged without screening applies, whenever credit risk is transferred via CDSs. Hence, the intermediating bank prefers to screen and use CCDSs to the outcome with CDSs, if and only if*

$$\mu(A^H - A^L) - \frac{C}{(1-\mu)} \geq (i - s(\mu))V_{0.2}^{fee}. \tag{10}$$

Whenever $(i - s^H)V_{0.2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H$*, the Screening Condition (10) reduces to* $(s_L - i)(1-\mu) \geq C/V_{0,2}^{fee}$ *as in the First Best.*

Proof. See the Appendix. □

Equation (10) results from a trade-off between the costs of overinvestment, the signaling premium, and screening costs. The outcome is a Second Best solution as shown in Table 2 ($SP > 0$). Whenever $(i - s^H)V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H$, the signaling premium is equal to zero ($SP = 0$). B then optimizes the trade-off between overinvestment and screening costs, yielding the same result as in the First Best. Again, market opacity does not impede the mechanism.

	Condition for screening	**Expected bank profit**	**Overall welfare**
First Best	$(1-\mu)(s^L - i) \geq C/V_{0,2}^{fee}$	$\mu(i - s^H)V_{0,2}^{fee}$ -C	$\mu(i - s^H)V_{0,2}^{fee}$ -C
CDS	No screening	$(i - s(\mu))V_{0,2}^{fee}$	$(i - s(\mu))V_{0,2}^{fee}$
CCDS ($SP > 0$)	$\mu(A^H - A^L) - (i - s(\mu))V_{0,2}^{fee}$ $\geq C/(1-\mu)$	$\mu(i - s^H)V_{0,2}^{fee}$ $-C - SP$	as in the First Best
CCDS ($SP = 0$)	as in the First Best	as in the First Best	as in the First Best

Table 2: A^H and A^L denote the fair value of the call feature on a high and a low recovery loan, respectively. C expresses screening costs, $V_{0,2}^{fee}$ is the present value of receiving one basis point of fee payments up to T_2 as long as there is no default, and SP is the signaling premium. The table summarizes results of credit risk transfer for high loan rates in four cases: (i) without information asymmetry (First Best), (ii) with information asymmetry using a CDS (CDS), (iii) with information asymmetry using a CCDS given a positive signaling premium (CCDS($SP > 0$)), (iv) with information asymmetry using a CCDS without a signaling premium (CCDS($SP = 0$)). For each case, the table outlines the condition for screening taking place (condition for screening), the expected bank profit given that the condition for screening is satisfied (expected bank profit), and the combined expected profit of the bank and the investor given the condition for screening is satisfied (overall welfare).

3.6 Discussion

Propositions 3 and 4 show that by structuring the hedge in a simple way, the basic intermediation dilemma can be solved: While credit derivatives permit a bank to relax binding lending constraints, a complete transfer of the underlying credit risk does not necessarily introduce costly disincentives with respect to screening. The First Best is still attainable with CCDSs.

The spread difference $(s^H - s^L)$ can be considered as a measure of information asymmetry. The term $(A^H - A^L)$ in Inequalities (8) and (10) is an increasing function of this difference, which leads to the following corollary:

Corollary 3.3. *Signaling with CCDSs is more preferable to standard credit risk transfer techniques the higher the information asymmetry.*

CCDSs can be important for financing over the life-cycle of a firm. A borrower may substantially grow and demand more funds after successful initial stage financing. A bank has usually acquired private information on the borrower during financing at early stages, which eases subsequent lending. At some point, the bank is, however, reluctant towards more loan grants to the same borrower due to concentration risk. A CCDS allows the bank to "grow" with the firm, because credit risk above a certain constraint can be fully transferred to a diversified investor without incurring information asymmetry costs. If continued loan grants intensify the relationship between the bank and the borrower, and support specialization, screening costs are reduced. Low screening costs, in turn, facilitate granting new loans.[19] CCDSs are, therefore, likely to enforce relationship building, credit portfolio management, and the rents generated therein.

The relevance of each loan rate scenario must be judged on a case by case basis as the banking structure, and its impact on loan rates charged by banks appear to differ across countries and sectors (Saunders and Schumacher (2000)). Maudos and Fernandez de Guevara (2004) obtain evidence for loan margin reductions in recent years. A trend towards lower loan rate scenarios stresses the importance of CCDSs as a mean to increase banks' profits and the overall welfare.

The high loan rate scenario addresses one of the central reasons for the recent credit crisis, namely irresponsible lending. Whereas, without securitization, credits are screened and only high loans are granted, this paper shows that standard credit risk transfer changes the incentive structure. In particular, a bank's profit is maximized by granting and transferring good and bad loans to the market *without* screening (overinvestment). The model, therefore, explains concerns of the IMF (2007) expressed in the Global Financial Stability Report about weakened credit discipline in conjunction with a strong credit volume growth. Relying on rating agencies alone does not seem to be a promising path to deal with the challenges of today's credit market.[20] I argue that the problems can instead be addressed by using appropriate credit risk transfer structures, which reestablish credit discipline, and, ultimately, solve the overinvestment problem.

For tractability, the model simplifies the pricing of credit risk transfer instruments. The results are, however, robust with respect to the notion that counterparty risk[21], random

[19]Tirole (2006) states that as it is costly for a borrower to find a new lender and for a bank to screen new borrowers, a natural market reaction to the existence of information asymmetries is to build up relationships between banks and borrowers.

[20]According to Tirole (2006), another layer of incentive problems arises with respect to the rating agency's behavior. Additionally, Ferreira and Schmidt (2006) argue that the rating of a rating agency is public information. It may not reveal the *specific* information needed to calculate the fair CDS rate.

[21]The investor may also default before the maturity of the contracts.

recovery rates, varying recovery payment conventions, the American feature of the call option, or a term structure of interest and loan rates can be introduced without affecting the basic signaling mechanism. The reason is that each of these considerations similarly affects the prices of credit risk transfer instruments on both high and low recovery loans, while the signaling mechanism simply depends on the difference between these prices.

3.7 The optimal contract

It remains to be analyzed whether a CCDS is the *optimal* security to solve the information asymmetry problem. A widely discussed alternative signaling mechanism is the partial risk retention contract. In contrast to CDSs and CCDSs, these contracts induce regulatory costs if bank capital is costly. The reason is that capital requirements are based on the maximum loss due to loan defaults in the Basel II jurisdiction.[22] In a partial risk retention contract, a bank retains a fraction θ of a loan's credit risk in order to signal its type. Suppose bank regulation requires to hold $\lambda > 0$ regulatory capital per unit of risk retained, and the unitary cost of capital is $\delta > 0$.[23] All types of risk retention contracts such as maturity mismatchs or first-to-default structures can be incorporated in this framework.

I assume that speculative banks participate in opaque credit derivative markets, and that μ is assigned to the probability of a fractional hedge θ being of high type. A bank with a high loan is then unable to credibly commit ex ante to retain a certain fraction if the costs of silently hedging the residual risk resulting from adverse beliefs, $V_{0,2}^{fee}(s(\mu) - s^H)\theta$, are lower than the regulatory costs of keeping the fractional risk, $\lambda\theta\delta$, i.e., if

$$\lambda\delta \geq V_{0,2}^{fee}(s(\mu) - s^H). \tag{11}$$

[22] Loan losses are completely covered by CDS and CCDS contracts. Hence, there are no capital requirements. In the standardized approach, the call needs to be at the discretion of the protection buying bank without positive incentives (for example a step-up in cost of cover) to call the protection before maturity (see BIS (2006) page 46 ff). Both conditions are met by the CCDS structure in my model. This treatment is justified from an economical point of view. Within risk retention contracts, such as maturity mismatches or first-to-default structures, the hedging costs may increase up to the time the hedge expires as a result of a deteriorating credit quality of the underlying. In contrast, the protection fee of a CCDS, at worst, remains constant up to maturity, justifying the full coverage assumption imposed.

[23] It is assumed that the unitary cost of capital is greater than the cost of deposits, which is normalized to zero. See, for example, Froot and Stein (1998).

Hence, partial risk retention contracts do not provide a credible solution to the incentive problem if regulatory capital requirements λ are high. CCDSs are the only viable signaling mechanism in this case.

The next proposition compares the expected bank profit with a CCDS to the one with a risk retention contract in the low loan rates scenario, assuming that regulatory costs are sufficiently low, i.e., $\lambda\delta \leq V_{0,2}^{fee}(s(\mu) - s^H)$.

Proposition 5. *A bank prefers a CCDS to a risk retention contract if*

$$\lambda\delta \geq \frac{((i - s^H)V_{0,2}^{fee} + C/(1-\mu) + A^L - A^H)(s^L - s^H)V_{0,2}^{fee}}{A^H - A^L}. \tag{12}$$

Proof. See the Appendix. □

Condition (12) shows that CCDSs are optimal if $A^H - A^L$ is sufficiently large, or if regulatory costs are high. Additionally, the term on the right side is directly linked to the costs of the signal in a CCDS contract, namely to $(i - s^H)V_{0,2}^{fee} + C/(1-\mu) + A^L - A^H$. This term is zero if CCDSs admit the First Best outcome according to Proposition 3. As $\lambda\delta > 0$ by the definition of regulatory costs, CCDSs are strictly preferred to risk retention contracts whenever the former induce the First Best in an environment of asymmetric information. Moreover, it can be shown that risk retention contracts never allow the First Best due to regulatory costs. Finally, signaling costs due to regulatory capital in a risk retention contract are wasteful, whereas the signaling premium of a CCDS in the Second Best directly accrues to the investor. He obtains a price for the call which more than compensates for the feature's economic value. Hence, in contrast to risk retention contracts, the overall welfare with CCDSs corresponds to the First Best whenever the screening condition is satisfied.

The analysis speaks to the current discussion on regulating the credit derivatives market. It is often argued that high regulatory capital is necessary to support financial stability. In contrast to this view, I show that high regulatory capital can prevent partial risk retention approaches, which have been extensively used before the recent credit crisis, from credibly signaling loan quality. My results suggest relying on properly structured credit risk transfer contracts such as CCDSs instead of turning to tighter regulatory restrictions.

4 Moral hazard

This section introduces monitoring and shows how moral hazard emerges in the model. It is beyond the scope of the paper to fully discuss the impact of debt contracts on monitoring within the bank-borrower relationship. For work in this field, see Innes (1990), Jensen and Meckling (1996), or Myers and Majluf (1984). Instead, I investigate the effect of hedging on monitoring.

Consider a bank B with the special ability to monitor a borrower F. Lacking a close relationship to F, the investor S is neither able to monitor nor to observe this activity.[24] He can, however, infer the bank's optimal monitoring-effort choice from the maximization problem. Coalition between B and F, and simultaneous defaults of S and F are excluded in the model.

4.1 Analysis

I start by analyzing moral hazard without adverse selection. The structure is as follows: At T_0, while granting a loan and transferring the credit risk, B fixes a monitoring-effort level e which is maintained up to T_2. The effort causes monitoring costs M(e) immediately payable at T_0. They are strictly increasing and convex, i.e., $M'(e) > 0$, $M''(e) > 0$. The protection leg (see the Appendix) is assumed to be strictly decreasing and convex, the fee leg to be strictly increasing and concave, and the fair CDS rate of a loan to be strictly decreasing and convex in monitoring effort, i.e., $s'(e) < 0$, $s''(e) > 0$.

A bank's optimal effort level is determined by the equality of marginal costs $M'(e)$ and marginal benefits of monitoring. An unhedged bank loses the fraction $1 - R$ on the nominal

[24] Without reputation effects and observability, the investor is also not able to influence the bank's maximization problem by paying for monitoring. In addition, Boot et al. (1993) show that the delegation of decision-making to the lender is optimal if the information revealed through monitoring is too detailed to contract upon.

in case of default, and the spread income $i - s(\cdot)$ after default. As B bears all benefits and costs of monitoring if it is not hedged, there is no moral hazard. In contrast, the incentives to monitor weaken for a CDS-hedged bank because it only loses the spread income in case of default.[25] Most of the marginal benefits of a hedged B's monitoring effort accrue to S, since the latter bears the risk of loan default. In this way, credit risk transfer originates moral hazard. A CCDS-structure provides partial relief for the problem under complete credit risk transfer. B's incentives to monitor the hedged position are enhanced, because monitoring rises the value of the embedded call feature. Hence, the possibility to call the contract internalizes part of the marginal benefits of monitoring. However, a bank's optimal effort level still remains below the one of an unhedged position.

Proposition 6. *B's incentives to monitor a loan hedged via CDS are small. A CCDS-hedge clearly induces more monitoring-effort, but still less than an unhedged loan position.*

Proof. See the Appendix. □

From an intuitive point of view, the result follows Innes' (1990) basic idea.[26] Choosing a CCDS instead of a monotonic CDS decreases a bank's obligation in "high" states (with a low CDS rate), and increases it in "low" states (with a high CDS rate): In high states, the bank can call the CCDS at T_1, resulting in a lower protection fee to be paid from T_1 to T_2. In low states, the additional premium for the call feature has been paid without a benefit at T_1. A higher monitoring-effort shifts probability weight to high states, implying that the CCDS contract is more likely to give the bank the opportunity to call at T_1. A larger benefit for B from monitoring-effort is created in this way. In respect thereof, a monotonic CCDS contract satisfies Innes' (1990) "maximal high-profit-state payoff" property, given the restriction that the credit risk of the loan must be fully transferred to an investor.

[25] The loss of spread income constitutes a risk as long as the face value of the defaulted loan can not be reinvested immediately at the same return.

[26] Innes (1990) proves the emergence of debt contracts in the presence of a monotonic contract constraint and the monotone likelihood ratio property (MLRP). A debt contract provides the best incentives for effort provision by extracting as much as possible from the entrepreneur under low performance, and by giving him the full marginal return from effort provision in high-performance states where revenues are above the face value of the debt.

4.2 Combination of adverse selection and moral hazard

To reach a plausible solution to the problem at hand, it is necessary to analyze whether the basic mechanism still maintains if adverse selection *and* moral hazard are present. As expected, the outcome is a combination of the previous results: A bank's maximization problem can be solved by relying on the pure adverse selection methodology, incorporating the influence of moral hazard on CDS rates.

Proposition 7. *The ability of CCDSs to induce beneficial screening is stronger if moral hazard and adverse selection are combined.*

Proof. See the Appendix. □

The rationale for Proposition 7 develops from the observation that CCDSs align the monitoring interests of B and S: A CCDS-hedge provides a certification of monitoring-effort. As this certification decreases the demanded protection fee, the conditions expressing whether CCDSs are beneficial are relaxed.

5 Conclusions

This paper explains why callable credit default swaps represent an optimal mechanism to transfer the credit risk of bank-originated loans. In particular, I model the credit risk transfer process from a bank to an investor. The standard transfer via CDSs causes information asymmetry costs because the bank can not credibly signal loan quality. I propose to use a callable credit default swap to address the problem. With this structure, a bank signals loan quality by expressing its readiness to pay for the implicit call feature. The ability to signal credit quality induces the bank to screen loans. Moreover, even though a CCDS completely transfers the loan's credit risk to an investor, the implicit call feature still encourages a bank to monitor the borrower. Hence, simple structuring avoids disincentives, and, consequently, reduces information asymmetry costs. It is further shown that the current capital requirements regulation can impede effective partial risk retention solutions. As a consequence, CCDSs evolve as the optimal credit risk transfer contract in most cases.

The findings speak to the debate on the impact of credit derivatives on intermediation. While the results support arguments on possible drawbacks of credit risk transfer due to disincentives, the broad implication of my analysis suggests that most of them can be approached by properly structuring credit derivatives. In particular, Duffee and Zhou (2001) conclude that the value of introducing a market for credit derivatives is ambiguous: Loan sales are initially used to transfer the risk of borrower default in both periods. One-period credit derivatives may simply replace loan sales, reducing the ability of banks to share the risk of borrower default in period two. In contrast, I show how banks can signal loan quality without incurring deadweight costs in period two, rendering the introduction of credit derivatives beneficial. Furthermore, the wealth destruction effect of a market for credit derivatives in Morrison (2005) - caused by a loss of the certification role of bank debt - can be avoided if banks do not cease to monitor the borrower. Debt hedged via CCDS indeed maintains its certification value in my model, because the bank keeps monitoring to increase the value of the call feature.

The model can be extended to include interactions between the bank and the borrower. For example, hold-up or collusion among these two players are likely to affect the outcome. It may be in the interests of a hedged bank and the borrower to delay default even though the recovery rate deteriorates by doing so, which harms the investor. A model incorporating the strategic game between all players promises further insights to what extent specialization, competition, and intermediation are affected by the new hedging tools.

References

George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 89:488–500, 1970.

Adam B. Ashcraft and João A. C. Sanots. Has the CDS market lowered the cost of corporate deb? *Journal of Monetary Economics*, 56:514–523, 2009.

Ross Barrett and John Ewan. BBA credit derivatives report 2006. *Britisch Bankers' Association*, 2006.

Sudipto Bhattacharya and Gabriella Chiesa. Proprietary information, financial intermediation, and research incentives. *Journal of Financial Intermediation*, 4:328–357, 1995.

BIS. International convergence of capital measurement and capital standards: A revised framework-comprehensive version. *Bank for International Settlements*, 2006.

Arnoud W. A. Boot, Stuart I. Greenbaum, and Anjan V. Thakor. Reputation and discretion in financial contracting. *American Economic Review*, 83:1165–1183, 1993.

Charles T. Carlstrom and Katherine A. Samolyk. Loan sales as a respnse to market-based capital constraints. *Journal of Banking and Finance*, 19:627–646, 1995.

Peter DeMarzo and Darrell Duffie. A liquidity-based model of security design. *Econometrica*, 67:65–99, 1999.

Douglas W. Diamond. Financial intermediation and delegated monitoring. *Review of Economic Studies*, 51:393–413, 1984.

Gregory R. Duffee and Chunsheng Zhou. Credit derivatives in banking: Useful tools for managing risks? *Journal of Monetary Economics*, 48:25–54, 2001.

Darrell Duffie. Innovations in credit risk transfer: Implications for financial stability. *Mimeo*, 2007.

Manuel Ferreira and Heinz Schmidt. *Was der Kapitalmarkt zur Bonität zu sagen hat.* 2006.

Günter Franke and Jan P. Krahnen. Default risk sharing between banks and markets: The contribution of collateralized loan obligations. *Mimeo*, 2005.

Günter Franke, Markus Herrmann, and Thomas Weber. Information asymmetries and securitization design. *Mimeo*, 2007.

Kenneth A. Froot and Jeremy C. Stein. Risk management, capital budgeting and capital structure policy for financial institution: An integrated approach. *Journal of Financial Economics*, 47:55–82, 1998.

Douglas Gale and Martin Hellwig. Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies*, 52:647–663, 1985.

William G. Gale. Federal lending and the market for credit. *Journal of Public Economics*, 42:177–193, 1990.

Gary Gorton and James A. Kahn. The design of bank loan contracts. *Review of Financial Studies*, 13:331–364, 2000.

Gary Gorton and George Pennacchi. Banks and loan sales, marketing nonmarketable assets. *Journal of Monetary Economics*, 35:389–411, 1995.

Greg M. Gupton, Christopher C. Finger, and Bhatia Mickey. Creditmetrics, the benchmark for understanding credit risk. *Technical Document, J.P. Morgan*, 1997.

Glenn H. Hubbard. Capital market imperfections and investment. *Journal of Economic Literature*, 36:193–225, 1998.

John Hull and Alan White. The valuation of credit default swap options. *Mimeo*, 2002.

IMF. Global financial stability report. September 2007.

Cho In-Koo and David M. Kreps. Signaling games and stable equilibria. *Quarterly Journal of Economics*, 102:179–222, 1987.

Robert Innes. Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, 52:45–67, 1990.

Michael C. Jensen and William H. Meckling. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3:305–360, 1996.

Benjamin J. Keys, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. Did securitization lead to lax screening? Evidence from subprime loans. *Mimeo*, 2008.

John Kiff and Ron Morrow. Credit derivatives. *Bank of Canada Review*, pages 3–11, 2000.

Joaqin Maudos and Juan Fernandez de Guevara. Factors explaining the interest margin in the banking sectors of the European Union. *Journal of Banking and Finance*, 28:2259–2281, 2004.

Alan D. Morrison. Credit derivatives, disintermediation and investment decisions. *Journal of Business*, 78:621–648, 2005.

Marek Musiela and Marek Rutkowski. *Martingale methods in financial modelling.* Springer, 2004.

Stewart C. Myers and Nicholas S. Majluf. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13: 187–221, 1984.

Antonio Nicolo and Loriana Pelizzon. Credit derivatives, capital requirements and opaque otc markets. *Mimeo*, 2008.

Dominic O'Kane, Marco Naldi, Sunita Ganapati, Arthur Berd, Claus Pedersen, Lutz Schloegl, and Roy Masal. *Guide to exotic credit derivatives.* Lehman Brothers, 2003.

J. Osborne, Martin. *An introduction to game theory.* Oxford University Press, Oxford, 2004.

Mitchell A. Petersen and Raghuram G. Rayan. The effect of credit market competition on lending relationships. *Quarterly Journal of Economics*, 110:407–443, 1995.

Timothy J. Riddiough. Optimal design and governance of asset-backed securities. *Journal of Financial Intermediation*, 6:121–152, 1997.

Bernard Salenié. *The economics of contracts.* MIT Press Cambridge, Massachusetts, 2005.

Anthony Saunders and Liliana Schumacher. The determinants of bank interest rate margins: An international study. *Journal of International Money and Finance*, 19:813–832, 2000.

Philipp J. Schönbucher. A tree implementaion of a credit spread model for credit derivatives. *Mimeo*, 1999a.

Philipp J. Schönbucher. A libor market model with default risk. *Mimeo*, 1999b.

Philipp J. Schönbucher. A note on survival measures and the pricing of options on credit default swaps. *Mimeo*, 2003.

Michael A. Spence. Job market signaling. *Quarterly Journal of Economics*, 87:355–374, 1973.

Joseph E. Stiglitz and Andrew Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71:393–410, 1981.

Jean Tirole. *The theory of corporate finance*. Princeton University Press 41 William Street, Princeton, New Jersey 08540, 2006.

Ernst-Ludwig Von Thadden. Long-term contracts, short-term investment and monitoring. *Review of Economic Studies*, 62:557–575, 1995.

Ernst-Ludwig Von Thadden. Asymmetric information, bank lending and implicit contracts: The Winner's Curse. *Mimeo*, 1998.

6 Appendix

6.1 Pricing

In this section, I define the basic valuation concepts for CDSs and options on CDSs. The fundamental references are Schönbucher (1999b), Schönbucher (2003), and Hull and White (2002). The pricing model is set up in a filtered probability space $(\Omega, (\mathcal{F}_T)_{(T \geq 0)}, \mathbb{P})$. Existence of a spot-martingale measure $\mathbb{Q} \sim \mathbb{P}$ is implied by the absence of arbitrage. The time of the credit event is modeled with a stopping time τ. I assume independence of default events, loan rates, and recovery rates. The riskless interest rate is zero.

6.1.1 The forward CDS rate

A CDS option's underlying is a forward starting credit default swap. A forward starting credit default swap is a CDS starting its life not immediately at T_0 but at T_1. If the reference entity defaults before T_1, the contract is worthless and no payments are made. The payoffs of a forward CDS can be split into a fee leg, paid by the protection buyer, and a protection leg, paid by the protection seller.

6.1.2 The fee leg

The protection buyer pays regular fee payments $s\delta_n 1_{T_n \leq \tau}$ at predetermined dates T_n to the protection seller. $1_{T_n \leq \tau}$ is the indicator function that default has not occurred before the payment date T_n. For tractability, I assume that defaults can only occur on one of the payment dates[27], and that the daycount fraction is equal to one. $V_{0.2}^{fee}$ denotes the value at time T_0 of receiving 1 basis point (bp) of fee payments up to T_2:

$$V_{0,2}^{fee} = E^Q[\sum_{n=0}^{N} 1_{T_n \leq \tau} \mid \mathcal{F}_{T_0}], \tag{13}$$

[27]Default on payment dates is a reasonable assumption given that defaults often become public at times when payments are due.

The value of the fee leg is, consequently, $sV_{0,2}^{fee}$. After a default, no fees are paid, i.e., $V_{\tau,2}^{fee} = 0$ for $\tau \leq T$. The forward starting fee leg is calculated in exactly the same way: I just adapt the summation in Equation (13) to include only the forward starting period. The resulting $V_{1,2}^{fee}$ is multiplied by the forward rate $f_{1,2}$ to arrive at the value of the forward starting fee leg at T_0: $f_{1,2}V_{1,2}^{fee}$.

6.1.3 The protection leg

In case of default during the life of the contract, the protection seller has to pay an amount 1-R to the protection buyer. Thus, the corresponding payment is $(1-R)1_{T_0 \leq \tau \leq T_2}$ at τ. The value of the protection leg at time T_0 is

$$V_{0,2}^{prot} = E^Q[(1-R)1_{T_0 \leq \tau \leq T_2} \mid \mathcal{F}_{T_0}]. \tag{14}$$

The value of the forward starting protection leg at T_0 is $V_{1,2}^{prot}$, obtained by replacing the indicator function in Equation (14) by $1_{T_1 \leq \tau \leq T_2}$ such that only the forward starting period is incorporated.

I now derive the CDS rate. The spot CDS rate simply equals the two legs. The fair *forward* CDS rate, $f_{1,2}$, at time T_0 of a forward CDS over the interval $[T_1, T_2]$ is the rate at which the forward fee leg has the same value as the forward protection leg, i.e.,[28]

$$f_{1,2} = \frac{V_{1,2}^{prot}}{V_{1,2}^{fee}}. \tag{15}$$

Markt-to-market values of income streams subject to credit risk can be calculated by multiplying with the corresponding fee leg. The value of a fixed spread received from T_0 to T_2, for example, is obtained by multiplying this spread by $V_{0,2}^{fee}$. Suppose a forward CDS over $[T_1, T_2]$ has been entered (as protection buyer) before or at T_0 at a rate x. The markt-to-market value of this position at time T_0 is $(f_{1,2} - x)V_{1,2}^{fee}$. The reason for this result is that one can lock into a fee stream of $(f_{1,2} - x)$ by selling the forward CDS protection at the current market rate $f_{1,2}$.

[28]Note that the fair forward CDS rate is not defined after default.

6.1.4 The change of numeraire technique

Take the example at the end of the previous paragraph. If, instead, one has just the right but not the obligation to enter the forward CDS at time T_0 at the forward CDS rate x, then the value of this right is[29]

$$(f_{1,2} - x)^+ V_{1,2}^{fee}. \tag{16}$$

Equation (16) corresponds to the payoff of an option on a forward-starting CDS with maturity T_0. Before deriving an option pricing formula, I need to introduce the change of numeraire technique.[30]

For any given spot-martingale-measure Q and numeraire A_T, one can define an equivalent pricing measure Q^A using the Radon-Nikodym density process:[31]

$$\frac{dQ^A}{dQ} |_T := \frac{A_T}{b_T} \frac{b_{T_0}}{A_{T_0}} \tag{17}$$

In our case the bank account is always 1, i.e., $b_{T_0} = b_T = 1$. Prices in the A-numeraire are martingales under the new measure Q^A.

Unfortunately, a direct application of V^{fee} as numeraire in a price system is not possible because V^{fee} can be zero. Thus, a price system in terms of V^{fee} is undefined after defaults.[32] Assume $\bar{A}_T$ to be the price process of a defaultable asset with zero recovery. For a given $T > T_0$, Schönbucher (2003) defines a "promised" payoff A'_T via $A'_T 1_{T<\tau} = \bar{A}_T$, which allows to define a default-free asset with the price process $E^Q[\frac{b_{T_0} A'_T}{b_T}]$. This asset pays A'_T at T for sure, i.e., the promised payoff is positive. In this case, it can be shown that $Q^{\bar{A}}$ is the measure which is reached when Q^A is conditioned on survival until T.[33] The new probability measure $Q^{\bar{A}}$ attaches zero probability to all events involving a default before T. Schönbucher (2003) circumvents the practical restrictions of this new interpretation by defining the value of XA'_T

[29] The reason is that one can now lock into a forward CDS contract at a rate $f_{1,2}$. The protection seller obtains a fixed spread $(f_{1,2} - x)$ over the future time interval $[T_1, T_2]$. It is shown in the last section why one simply needs to multiply a spread over this future period by $V_{1,2}^{fee}$ to obtain its value.

[30] See Schönbucher (2003).

[31] See, for example, Musiela and Rutkowski (2004).

[32] While the price system itself can break down if the numeraire is zero, the Radon-Nikodym density still remains valid as long as $A_{T_0} > 0$. It defines a mathematically admissible Radon-Nikodym density.

[33] Q^A is the measure for a default-free numeraire defined via the Radon-Nikodym density process.

at T in survival:

$$E^Q[\frac{b_{T_0}}{b_T}1_{\tau>T}XA'_T \mid \mathcal{F}_t] = \bar{A}_{T_0}E^{Q^{\bar{A}}}[X \mid \mathcal{F}_{T_0}] \tag{18}$$

The right hand side of Equality (18) is obtained by plugging in the Radon-Nikodym density process with the defaultable asset as numeraire.

6.1.5 Option pricing formula

Armed with the survival-based pricing measure I use the defaultable fee stream for a change of measure to a survival measure. Equation (18) is applied to value a receiver swaption on $f_{1,2}$ at T_0 with maturity T_1 and strike x:

$$RS(T_0) = 1_{\tau>T_0}E^Q\big(\frac{b_{T_0}}{b_{T_1}}(x - f_{1,2}(T_1))^+V_{1,2}^{fee}(T_1)1_{\tau>T_1}|\mathcal{F}_{T_0}\big) \tag{19}$$

$$= 1_{\tau>T_0}V_{1,2}^{fee}E^{P^{V_{1,2}^{fee}}}\big((x - f_{1,2}(T_1))^+ \mid \mathcal{F}_{T_0}\big) \tag{20}$$

The expectation is taken under the $P^{V_{1,2}^{fee}}$-measure. T_1 in parenthesis indicates that T_1 is the starting date. The key is to understand that $V_{1,2}^{fee}(T_1)1_{(\tau>T_1)}$ in Equation (19) corresponds to the price process of a defaultable asset with zero recovery at time T_1. We can, therefore, plug in the corresponding Radon-Nikodym density to generate Equation (20). This point is the furthest one can go without making any modeling assumptions apart from the absence of arbitrage. For examples of how to model the distribution of the forward CDS rate $f_{1,2}(T_1)$ under the P^{Vfee}-measure, see Schönbucher (2003).

It is important to realize that the underlying credit risk assessment is not necessary to *price* the option itself. This assessment has already been done when $f_{1,2}$ was formed.

6.2 Proofs

Proposition 1. *For a low loan rate, the bank screens a new loan applicant in the First Best as long as* $\mu(i - s^H) \geq C/V_{0,2}^{fee}$.

Proof. $V_{0,2}^{fee}$ is used to convert risky spread-incomes from T_0 to T_2 into present values. A bank selects a strategy by comparing expected profits. When doing so, it anticipates the protection fee for each strategy on the credit derivatives market. The bank's expected profit from screening and subsequent decision about loan granting is $\mu V_{0,2}^{fee}(i - s^H) - C$. A detected low credit is not approved, since i$<s^L$. Just rejecting a loan without screening the credit quality yields zero expected profit, and granting a loan without screening gives $V_{0,2}^{fee}(i - s(\mu)) < 0$. Comparing expected profits and rearranging terms yields the condition in Proposition 1. $\square$

Proposition 2. *Information asymmetry between the bank B and the investor S impedes screening activities, if lending constraints are relaxed by transferring the credit risk via CDSs. The loan market breaks down, resulting in underinvestment.*

Proof. The Intuitive Criterion of In-Koo and Kreps (1987) can be informally explained as follows: Suppose B makes an out-of-equilibrium offer. B does so based on a certain conjecture about how the investor reacts. The Intuitive Criterion suggests to assign probability one to a B having a high recovery loan, if and only if given the investor's most optimistic conjecture, a high bank finds it optimal to make this offer and to deviate from the equilibrium, while the low bank does not.

Selling a risk on the CDS market permits no credible quality signal. Hence, it is easily checked that the only equilibrium surviving the Intuitive Criterion is a pooling equilibrium without market activity. The bank uses a CDS to hedge the exposure, and S assumes that no screening is taking place. The latter only accepts a protection fee equal or higher than $s(\mu)$, no matter what B claims about the credit's quality. Spending C in this equilibrium reduces B's expected profit: Screening leads to a NPV of Max[0,$V_{0,2}^{fee}(i - s(\mu))] - C \leq Max[0, V_{0,2}^{fee}(i - s(\mu))]$. The bank is better off if it drops screening and stays out of the market, since spending C does not allow to reduce the protection fee.

A necessary condition for the existence of a pooling equilibrium without market activity is that B makes negative profits by offering any contract. This condition is satisfied because the loan rate is smaller than $s(\mu)$. Hence, the expected profit is maximized if B stays out of the market. The market breaks down, which causes underinvestment since loans with positive NPVs are also rejected. □

Lemma 3.2. *The value of a receiver swaption is increasing in the recovery rate:*

$$\frac{\partial RS}{\partial R} > 0. \tag{21}$$

Proof. a) Changing R in Equation (20) does neither alter $V_{1,2}^{fee}$ nor $V_{0,2}^{fee}$. Consequently, the Radon-Nikodym density remains unchanged, and I do not need to take a changing survival-based pricing measure into account. $f_{1,2}(T_1) = \frac{V_{1,2}^{prot}(T_1)}{V_{1,2}^{fee}(T_1)}$ can be written as $\frac{(1-R)V_{1,2}^{Fprot}(T_1)}{V_{1,2}^{fee}(T_1)} = (1-R)f_{1,2}^F$, where V^{Fprot} is the value of a claim paying 1 in case of default. Obviously, using Equation (20), $RS(T_0) = 1_{\tau>T_0}V_{1,2}^{fee}E^{P^{V_{1,2}^{fee}}}\left((x - f_{1,2}(T_1))^+ \mid F_{T_0}\right) = 1_{\tau>T_0}V_{1,2}^{fee}E^{P^{V_{1,2}^{fee}}}\left((x - (1-R)f_{1,2}^F(T_1))^+ \mid F_{T_0}\right) = 1_{\tau>T_0}V_{1,2}^{fee}\int_0^{x/(1-R)}(x-(1-R)f_{1,2}^F\varphi(f_{1,2}^F)\partial f_{1,2}^F$, with φ expressing the density of $f_{1,2}^F$. Now define F(Z,R) $= \int_0^Z(x-(1-R)f_{1,2}^F\varphi(f_{1,2}^F)\partial f_{1,2}^F$. Then, $\frac{\partial}{\partial R}F(\frac{x}{(1-R)},R) = \frac{\partial}{\partial Z}F(Z,R)\mid_{Z=\frac{x}{1-R}}$ $+\frac{\partial}{\partial R}F = \int_0^Z f_{1,2}^F\varphi(f_{1,2}^F)\partial f_{1,2}^F > 0$, since $\frac{\partial}{\partial Z}F(Z,R)\mid_{Z=\frac{x}{1-R}} = 0$, $f_{1,2}^F > 0$ and $\varphi(f_{1,2}^F) > 0$. As $1_{\tau>T_0}$ and $V_{1,2}^{fee}$ are independent of R, it follows immediately that the RS is increasing in R. □

Proposition 3. *The intermediating bank prefers to screen and use CCDSs to the outcome with CDSs, if and only if*

$$\mu(A^H - A^L) - \frac{C}{(1-\mu)} \geq 0. \tag{22}$$

Whenever $(i-s^H)V_{0,2}^{fee}+A^L+\frac{C}{(1-\mu)} \leq A^H$, *the Screening Condition (8) reduces to* $\mu(i-s^H) \geq C/V_{0,2}^{fee}$ *as in the First Best.*

Proof. Consider a CCDS-contract that is only callable once at T_1, and that the true recovery rate over the remaining time period to maturity T_2 is publicly available once this point in time is reached. B tries to maximize expected profits in the set of incentive-compatible, profitable-type-by-type contracts. Assume that B chooses between as many contractual terms as there are possible types of loans. According to Salenié (2005), the revelation principle then implies that one can restrict attention to mechanisms that are both direct (where B reports his information) and truthful (so that B finds it optimal to announce the true value of his information). In other words, the optimal contract induces B to reveal his type. The purpose of contract choice is to influence beliefs. To maximize the expected profits of the loan and the credit derivative contract by signaling quality, a bank detecting a high recovery loan solves the following program:

$$\max_{P}\big(iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - A^{H}) - C\big) \tag{23}$$

s.t.

$$P \geq A(\rho) \;\; (IR) \tag{24}$$

$$(i - s(\rho))V_{0,2}^{fee} + A^{L} + \frac{C}{(1-\mu)} \leq P \;\; (IC) \tag{25}$$

This set of equations is justified by the following procedure to obtain a perfect Bayesian equilibrium in pure strategies: The bank chooses the swaption premium$P$ by anticipating the resulting protection fee on the credit derivatives market. S signs the credit risk transfer contract, demanding a protection fee dependent on beliefs ρ consistent with the bank's strategy. The contract is set in a way to maintain $\rho = 1$.

Equation (23) shows the expected profit of a bank with a high recovery loan. Equation (24) is the participation constraint (IR) of S if he assumes that the bank offering a CCDS is indeed of high type with probability ρ. S, at least, demands a premium $A(\rho)$ for the call feature. The incentive compatibility constraint (IC) expresses two conditions: A bank with a low recovery credit stays out of the market after screening: $iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - A^{L}) \leq 0$. More restricting, a bank must also weakly prefer to "screen and choose an adequate contract" to just pick the CCDS contract without screening at all:[34] $iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - \mu A^{H} - (1-\mu)A^{L}) \leq \mu\big(iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - A^{H}) - C$. This condition is necessary because S cannot observe the screening-activity itself. The lowest premium P satisfying the IC and the IR is labeled P*.

[34] The weak monotonic profit assumption allows to skip the IC of a bank with a high recovery loan. See, for example, Tirole (2006).

P* guarantees equilibrium beliefs being such that $\rho = 1$ for a CCDS offer if and only if (i) it is profitable to offer P* after detection of a high credit, (ii) it is profitable to screen. Since $(ii) \Rightarrow (i)$, and the alternative outcome is a market breakdown, this condition looks as follows:

$$\mu(iV_{0,2}^{fee} - (P^* + s^H V_{0,2}^{fee} - A^H)) - C + (1-\mu)0 \geq 0. \tag{26}$$

P* induces $s(\rho) = s^H$ and $A(\rho) = A^H$. The right hand side of Equation (26) corresponds to the expected profit in a market breakdown. The left hand side expresses the choice of a CCDS for a high loan, and staying out of the market for a low loan after screening. Public information implies that - before screening - B assumes to be able to offer a CCDS contract with probability μ. If Inequality (26) is satisfied for P*, a separating equilibrium with screening holds: A bank with a high loan chooses a CCDS, and staying out of the market is the only rational response after the detection of a low recovery credit. The CCDS (i.e., P* in the IC) is too expensive for a B with a low loan, and offering a plain vanilla CDS entails a bad signal.[35]

Rearranging (26) and plugging in P^* from the IC yields

$$\mu(A^H - A^L) - \frac{C}{(1-\mu)} \geq 0. \tag{27}$$

The resulting separating equilibrium satisfies the Intuitive Criterion. To check uniqueness, note that a bank with a high recovery loan chooses the minimum swaption premium still allowing to signal its type. Consequently, the least-cost separating equilibrium must leave a B with a low loan just indifferent between the two contracts. It is the most efficient separating perfect Bayesian equilibrium for B, as it entails the lowest signaling premium transferred to S.

[35]The reason is that by observing the public known parameters, S infers that screening is optimal. Hence, - using the Intuitive Criterion - only a bank with a low recovery loan subsequently tries to offer a CDS. According to (i) and (ii), a bank with a high recovery loan prefers a CCDS. The updated beliefs of S for a CDS-offer are $\rho = 0$, yielding a protection fee s^L above the loan rate. Naturally, B prefers to stay out of the market than paying this protection fee in a CDS offer.

Now, consider the following case: If

$$(i - s^H)V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H, \tag{28}$$

then the bank with a high recovery loan only needs to pay A^H for the call feature in order to satisfy the IR and the IC. Hence, one concludes that $P^* = A^H$. Plugging this premium into Equation ((26)) to check the conditions for a screening equilibrium, one obtains $\mu(i - s^H) \geq C/V_{0,2}^{fee}$. There is no difference to the screening condition in the First Best, and $P^* = A^H$ also induces the First Best expected profits.

Note that a bank may silently transfer the call feature in period 2 by addi- tionally *selling* a receiver swaption. It is easily shown why this possibility does not affect the result: Whenever it is optimal for a high bank to sell a RS, it is also optimal to do so for a low bank. As long as $\mu < 1$, the high bank is, consequently, faced with unfavorable beliefs and obtains a price below A^H if it sells the right to call. Hence, selling a RS would decrease expected profits. As a result, the IC is not tightened by the possibility of silent credit risk transfer.

Suppose next that it is not publicly known whether the bank's credit risk constraint is binding: Speculative banks without lending restrictions may also try to sell receiver swaptions. Assuming most unfavorable beliefs of the investor, i.e., the bank being speculative with probability 1, a high B needs to satisfy the following IC:

$$(i - s(\rho))V_{0,2}^{fee} + A(\mu) + \frac{C}{(1-\mu)} \leq P \ \ (IC) \tag{29}$$

Inequality (29) expresses that a low bank silently selling a receiver swaption obtains $A(\mu)$ for this right, as investors assume it is a speculative bank.[36] The new screening condition is easily derived as

$$\mu(A^H - A(\mu)) - \frac{C}{(1-\mu)} \geq 0. \tag{30}$$

The information about whether the lending constraint is binding is not necessary to maintain

[36] A speculative bank without binding lending constraint will not pay the high premium P^*. However, the possibility of a speculative bank selling a RS has the following effect: A low bank can pretend to be speculative and, hence, sell the RS at a higher price.

a separating equilibrium. $\square$

Proposition 4. *Consider a market characterized by a high loan rate. A pooling equilibrium with loans being hedged without screening applies, whenever credit risk is transferred via CDSs. Hence, the intermediating bank prefers to screen and use CCDSs to the outcome with CDSs, if and only if*

$$\mu(A^H - A^L) - \frac{C}{(1-\mu)} \geq (i - s(\mu))V_{0.2}^{fee}. \tag{31}$$

Whenever $(i - s^H)V_{0.2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H$*, the Screening Condition (31) reduces to* $(s_L - i)(1-\mu) \geq C/V_{0,2}^{fee}$ *as in the First Best.*

Proof. The program to maximize takes the following form:

$$\max_P \big(iV_{0,2}^{fee} - (P + s(\rho)V_{0,2}^{fee} - A^H) - C\big) \tag{32}$$

s.t.

$$P \geq A(\rho) \tag{33}$$

$$(i - s(\rho))V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq P \tag{34}$$

The first part of the proof corresponds to the one in Proposition 3. The second part is slightly different. Instead of Inequality (7), I need to check if

$$\mu(iV_{0,2}^{fee} - (P^* + s^H V_{0,2}^{fee} - A^H)) - C + (1-\mu)0 \geq iV_{0,2}^{fee} - (\mu s^H + (1-\mu)s^L)V_{0,2}^{fee}, \tag{35}$$

since the CDS-outcome is a pooling equilibrium with market activity. Just using CDSs is the better choice for B if Inequality (35) is not satisfied. Plugging in P^* from Equation ((34)) yields

$$\mu(A^H - A^L) - \frac{C}{(1-\mu)} \geq (i - s(\mu))V_{0,2}^{fee}. \tag{36}$$

Additionally, one can replace P* by A^H in Equation (35) if $(i - s^H)V_{0,2}^{fee} + A^L + \frac{C}{(1-\mu)} \leq A^H$. Simplifying gives the same screening condition as in the First Best.

Using the same argument as in the proof of Proposition 3, privacy of credit derivative contracts does not change the results. Additionally, allowing for speculative banks slightly affects the outcome. In the worst case, i.e., if the investor assumes the bank to sell a RS for speculative reasons with probability 1, Condition (36) looks as follows:

$$\mu(A^H - A(\mu)) - \frac{C}{(1-\mu)} \geq (i - s(\mu))V_{0,2}^{fee} \tag{37}$$

□

Proposition 5. *A bank prefers a CCDS to a risk retention contract if*

$$\lambda\delta \geq \frac{((i - s^H)V_{0,2}^{fee} + C/(1-\mu) + A^L - A^H)(s^L - s^H)V_{0,2}^{fee}}{A^H - A^L}. \tag{38}$$

Proof. Consider the risk retention contract. A bank detecting a high recovery loan solves the following program:

$$\max_{\theta}\left(iV_{0,2}^{fee} - \theta s^H V_{0,2}^{fee} - (1-\theta)s(\rho)V_{0,2}^{fee} - C - \delta\lambda\theta\right) \tag{39}$$

s.t.

$$\frac{(i - s^H)V_{0,2}^{fee} + \frac{C}{(1-\mu)}}{\delta\lambda + V_{0,2}^{fee}(s^L - s^H)} \leq \theta \tag{40}$$

The participation constraint is already incorporated in Equation (39), i.e., a competitive credit derivative market requires a fair premium to the investor. Inequality (40) expresses the incentive compatibility constraint: A low bank or a bank which does not screen incurs a loss if it retains θ. The high bank will choose the smallest possible fraction θ^* satisfying Inequality (40), as retaining risk is costly due to the required regulatory capital.

Finally, the signaling costs of a CCDS must be lower than the signaling costs of a risk retention contract, i.e., $P^* - A^H \leq \delta\lambda\theta^*$, which is expressed in Condition (12). □

Proposition 6. *B's incentives to monitor a loan hedged via CDS are small. A CCDS-hedge clearly induces more monitoring-effort, but still less than an unhedged loan position.*

Proof. For an unhedged, granted loan the situation looks as follows: B earns a risky loan rate i, and bears the corresponding risk of a default. The first part has a value of $iV_{0,2}^{fee}$ at T_0. The second part is simply the protection leg. It is known that the CDS rate at time T_0 is the rate at which the fee leg has the same value as the protection leg. I, therefore, express the latter's value as $s^{H,L}V_{0,2}^{fee}$. The bank's problem can be treated separately for each loan type. Hence, the superscript (H,L) is suppressed in what follows. B bears the costs and full benefits of the monitoring activities. The IR is binding and, hence, directly included. The optimal monitoring-effort level is derived by solving the following maximization problem:

$$\max_e \big((i - s(e))V_{0,2}^{fee}(e) - M(e)\big) \tag{41}$$

Optimization equates B's marginal benefits and marginal costs of effort:

$$(i - s(e))V_{0,2}^{\prime fee}(e) - s'(e)V_{0,2}^{fee}(e) = M'(e) \tag{42}$$

The effort level e solving (42) is efficient as it maximizes the lender's value from monitoring. It represents the First Best level of effort.

If one considers a position hedged via CDS, the bank maximizes the following expression:

$$\max_e \big((i - s)V_{0,2}^{fee}(e) - M(e)\big) \tag{43}$$

Maximization yields

$$(i - s)V_{0,2}^{\prime fee}(e) = M'(e). \tag{44}$$

The effort level $\hat{e}$ solving (44) is strictly lower than e, because $s'(e) < 0$, $V_{0,2}^{fee}(e) \geq 0$, and $M(e)$ is convex per assumption. Without reinvestment risk, $\hat{e}$ is even equal to zero. It remains to be shown why CCDSs yield an effort level in between e and $\hat{e}$. Optimizing

$$\max_e \big((i - s)V_{0,2}^{fee}(e) - P + A(e) - M\big) \tag{45}$$

yields

$$(i-s)V'^{fee}_{0,2}(e) + A'(e) = M'(e). \tag{46}$$

To show that $\frac{\partial RS(T_0)}{\partial e} > 0$, i.e., $A'(e) > 0$, recall the assumption that effort is maintained up to T_2, $s'(e) < 0$, $f'_{1,2}(e) < 0$, $\frac{\partial V^{prot}_{1,2}}{\partial e} < 0$, and $\frac{\partial V^{fee}_{1,2}}{\partial e} > 0$. Using the Radon-Nikodym density, the swaption pricing formula (20) can be expressed in terms of the probability measure Q, which is unaffected by a change of e: $1_{\tau>T_0}E^Q\big(1_{\tau>T_1}(x-f_{1,2}(T_1))^+V^{fee}_{1,2}(T_1) \mid \digamma_{T_0}\big)$. It remains to be shown that $\frac{\partial 1_{\tau>T_0}E^Q\left(1_{\tau>T_1}(x-f_{1,2}(T_1))^+V^{fee}_{1,2}(T_1)|\digamma_{T_0}\right)}{\partial e} > 0$. The latter expression can be written as $1_{\tau>T_0}\frac{\partial}{\partial e}E^Q[g(e,Y)h(e,Y)1_{min(k(e,Y),m(e,y))}] > 0$. $Y := (y\epsilon\mathbf{R} \mid 1_{min(k(e,Y),m(e,y))>0})$. If k and m are continuous and differentiable functions, and $\varepsilon > 0$ but small, define $e_2 := e_1 + \varepsilon$. As $\frac{\partial k}{\partial e} > 0$ and $\frac{\partial m}{\partial e} > 0$, I have $k(e_2,y_1) > k(e_1,y_1)$ and $m(e_2,y_1) > m(e_1,y_1) \Rightarrow min(k(e_2,y_1),m(e_2,y_1)) > min(k(e_1,y_1),m(e_1,y_1))$. Consequently, if $Y_e = (a_e,b_e)$ and $y\epsilon Y_e$, one concludes that $(a_{e1},b_{e1}) < (a_{e2},b_{e2})$ $\forall e$. Hence, $1_{\tau>T_0}\frac{\partial}{\partial e}E^Q[\big(g(e,Y)h(e,Y)1_{min(k(e,Y),m(e,y))}]= \lim_{h\to 0}E^Q[\big(g(e+h,Y)h(e+h,Y)1_{y\epsilon(a_{e+h},b_{e+h})} - g(e,Y)h(e,Y)1_{y\epsilon(a_e,b_e)}\big)\frac{1}{h}]=E^Q[\lim_{h\to 0}\frac{1}{h}\big(g(e+h,Y)h(e+h,Y)1_{y\epsilon(a_{e+h},b_{e+h})}-g(e,Y)h(e,Y)1_{y\epsilon(a_e,b_e)}\big)]$, given that $g(e,Y)h(e,Y)1_{y\epsilon(a_e,b_e)}$ is assumed to be evenly integrable.[37] The latter term is equal to $E^Q[\lim_{h\to 0}\frac{1}{h}(g(e+h,Y)h(e+h,Y)1_{y\epsilon(a_e,b_e)} - g(e,Y)h(e,Y)1_{y\epsilon(a_e,b_e)})]$, using the previous result $(a_{e1},b_{e1}) < (a_{e2},b_{e2})$. Rearranging terms gives $E^Q[1_{y\epsilon(a_e,b_e)}\lim_{h\to 0}\frac{1}{h}\big(g(e+h,Y)h(e+h,Y) - g(e,Y)h(e,Y)\big)]$, which is, by definition, equal to $E^Q[1_{y\epsilon(a_e,b_e)}\big((\frac{\partial}{\partial e}g(e,Y))h(e,Y) + g(e,Y)(\frac{\partial}{\partial e}h(e,Y))\big)]$. Consequently, the value of the RS is increasing in effort, as each term within the brackets of the latter equation is greater than zero.

The preceding result leads to a strictly higher optimal monitoring-effort level $\bar{e}$ in ((46)) on a CCDS hedged loan than on a CDS hedged credit. Finally, arbitrage principles imply that $A'(e) < -s'(e)V^{fee}_{0,2}(e)$: This inequality follows because $-s'(e)V^{fee}_{0,2}(e)$ corresponds to $-V'^{prot}_{0,2}(e)$, and $-V'^{prot}_{0,2}(e) > -V'^{prot}_{1,2}(e)$. By the definition of conditional claims, $A'(e)$ must be smaller or equal to $-V'^{prot}_{1,2}(e)$. A receiver swaption is only a conditional claim on the change in the value of the forward protection leg. Consequently, I have a higher monitoring-effort on an unhedged loan than on a CCDS hedged position.[38] □

[37] It is sufficient to show the existence of a integrable majorable.

[38] One can certainly attain the same level of monitoring if the bank leverages the CCDS position by buying

Proposition 7. *The ability of CCDSs to induce beneficial screening is stronger if moral hazard and adverse selection are combined.*

Proof. Denote by e the First Best effort level, by $\bar{e}$ the optimal effort level for a CCDS hedged loan, and by $\hat{e}$ the one for a CDS hedged credit. As shown in Proposition 6, $\hat{e}$ can be assumed to be equal to zero without loss of generality. The credit spread $s(e)$ reflects the protection fee given S supposes the First Best monitoring effort level. Similarly, s($\bar{e}$) and $s(\hat{e})$ are the protection fees payable for a CCDS and a CDS hedge, respectively. According to Proposition 6, $s(\hat{e}) > s(\bar{e}) > s(e)$, as the investor anticipates the bank's monitoring incentives. The corresponding monitoring costs are $M(e) > M(\hat{e}) > M(\hat{e}) = 0$. In the same way, the notations $V_{0,2}^{fee}(\cdot)$ and $A(\rho, \cdot)$ indicate that these values are also affected by monitoring.

The case of high interest rates is analyzed first: B solves the following program:

$$\underset{P}{max}\big(iV_{0,2}^{fee}(\bar{e}) - (P + s(\rho, \bar{e})V_{0,2}^{fee}(\bar{e}) - A^H(\bar{e})) - C - M(\bar{e})\big) \tag{47}$$

s.t.

$$P \geq A(\rho, \bar{e}) \tag{48}$$

$$(i - s(\rho, \bar{e}))V_{0,2}^{fee}(\bar{e}) + A^L(\bar{e}) + \frac{C}{(1-\mu)} - M(\bar{e}) \leq P \tag{49}$$

The IC, i.e., Inequality (49), is simplified: I do not account for the fact that $A'(\bar{e})$ is different for a high, respectively a low loan. In other words, using Equation (46), the monitoring-effort level is assumed to be the same for a high and a low loan hedged via CCDS at the same protection fee.

To incorporate moral hazard into the optimal decision of the bank, I rewrite Condition (35):

$$\mu(iV_{0,2}^{fee}(\bar{e}) - (P^* + s^H(\bar{e})V_{0,2}^{fee}(\bar{e}) - A^H(\bar{e})) - M(\bar{e})) - C + (1-\mu)0$$

more RSs than necessary to unwind the position at T_1. However, in this case the full insurance restriction is not satisfied. As in Innes (1990), there is always underprovision of effort relative to the First Best, where the bank does not need to hedge.

$$\geq iV_{0,2}^{fee}(\hat{e}) - (\mu s^H(\hat{e}) + (1-\mu)s^L(\hat{e}))V_{0,2}^{fee}(\hat{e}) - M(\hat{e}) \tag{50}$$

The right hand side of Inequality (50) incorporates the following effects: The optimal effort level for the CDS-hedge alternative is $\hat{e} < \bar{e} < e$. Hence, monitoring costs decrease to $M(\hat{e})$. However, lower effort is anticipated by S which increases the protection fee, and decreases the valuation of fee streams.

Rearranging Condition (50) and plugging in P^* yields

$$\mu(A^H(\bar{e}) - A^L(\bar{e})) - \frac{C}{(1-\mu)} \geq (i - s(\mu,\hat{e}))V_{0,2}^{fee}(\hat{e}) - M(\hat{e}) = (i - s(\mu))V_{0,2}^{fee}. \tag{51}$$

Compared to the adverse selection case without the possibility of monitoring, the screening condition is unchanged. The reason is that the right hand side of Inequality (51) is the same as in (31), because $\hat{e} = 0$. The left hand side also remains unchanged.

If

$$(i - s(\rho,\bar{e}))V_{0,2}^{fee}(\bar{e}) + A^L(\bar{e}) + \frac{C}{(1-\mu)} - M(\bar{e}) \leq A(\rho,\bar{e}), \tag{52}$$

then B only pays $A(\rho,\bar{e})$ for the call feature. Plugging this premium into Inequality (50) for P^* yields

$$\mu\big((i - s^H(\bar{e}))V_{0,2}^{fee}(\bar{e}) - M(\bar{e})\big) - C \geq (i - s(\mu,\hat{e}))V_{0,2}^{fee}(\hat{e}) - M(\hat{e}). \tag{53}$$

Equalities (42) and (46) imply $(i - s^H(\bar{e}))V_{0,2}'^{fee}(\bar{e}) - s^{H\prime}(\bar{e})V_{0,2}^{fee}(\bar{e}) > M'(\bar{e})$. Consequently, $(i - s^H(\bar{e}))V_{0,2}^{fee}(\bar{e}) - M(\bar{e}) - \big((i - s^H(0))V_{0,2}^{fee}(0) - M(0)\big) = D > 0$, and I can rewrite (54) as follows:

$$\mu\big((i - s^H)V_{0,2}^{fee} + D\big) - C \geq (i - s(\mu)V_{0,2}^{fee} \tag{54}$$

Simplifying and rearranging terms to $(s^L - i)(1-\mu) \geq (C - \mu D)/V_{0,2}^{fee}$ shows that the screening condition is relaxed compared to Proposition 4.

In the case of low interest rates, one needs to check whether

$$\mu(iV_{0,2}^{fee}(\bar{e}) - (P^* + s^H(\bar{e})V_{0,2}^{fee}(\bar{e}) - A^H(\bar{e})) - M(\bar{e})) - C + (1-\mu)0 \geq 0. \tag{55}$$

After replacing P^*, Inequality (55) yields the following, unchanged screening condition:

$$\mu(A^H(\bar{e}) - A^L(\bar{e})) - \frac{C}{(1-\mu)} \geq 0 \tag{56}$$

If Inequality (52) is satisfied, the following holds: $\mu\big((i - s^H(\bar{e}))V_{0,2}^{fee}(\bar{e}) - M(\bar{e})\big) \geq C$, or $\mu(i - s^H) \geq (C - \mu D)/V_{0,2}^{fee}$. The latter condition is relaxed compared to the case without monitoring in Proposition 3. □

Macroeconomic Conditions, Growth Options and the Cross-Section of Credit Risk

Marc Arnold, Alexander F. Wagner, Ramona Westermann *

June 20, 2011

ABSTRACT

This paper develops a structural equilibrium model with intertemporal macroeconomic risk, incorporating the fact that firms are heterogeneous in their asset composition. Compared to firms which are mainly composed of invested assets, firms with growth options have larger costs of debt because they are more volatile and have a higher tendency to default during recession when marginal utility is high and recovery rates are low. Our model matches stylized facts regarding credit spreads, default probabilities, leverage, and investment clustering. Importantly, it also makes predictions about the cross-section of all these features.

JEL-code: G32

Keywords: Capital structure, macroeconomic risk, growth options, credit spread puzzle

*Arnold is a Ph.D. student at the Swiss Finance Institute at the University of Zurich, and Westermann a Ph.D. student at the Swiss Finance Institute at the University of Geneva. Wagner is an assistant professor of finance at the Swiss Finance Institute at the University of Zurich and a CEPR Research Affiliate. Office address: Department of Banking and Finance, Plattenstrasse 14, CH-8032 Zurich, Switzerland, Phone: +41-44-634-3963, Email: alexander.wagner@bf.uzh.ch. We thank an anonymous referee whose suggestions have greatly improved the paper. Tony Berrada, Simon Broda, Marc Chesney, Pierre Collin-Dufresne, Rajna Gibson, Michel Habib, Dirk Hackbarth, Mario Häfeli, Erwan Morellec, Gabriel Neukomm, Kjell Nyborg, Tatjana-Xenia Puhan, Alexandre Ziegler and seminar participants at the University of Zurich, the Humboldt University of Berlin, the C.R.E.D.I.T Conference in Venice, and the Research Day of the NCCR FINRISK provided helpful comments. This research was supported by the NCCR FINRISK, the Swiss Finance Institute, and the Research Priority Program "Finance and Financial Markets" of the University of Zurich.

1. Introduction

This paper examines the effect of time-varying macroeconomic conditions on credit spreads, firm value, and financial policy choices of firms with assets in place and growth options. The central thesis we develop is that expansion options react quite differently to the underlying risk in the economy than invested assets do. First, as growth options are levered claims, they are more volatile than assets in place. Second, we show that firms with lots of growth options ("growth firms") are more exposed to business cycle risk, which induces a stronger tendency to default during recession when recovery rates are low and marginal utilities high. Both effects raise the costs of debt and induce lower leverage choices of growth firms compared to those of firms with mainly invested assets.

Standard structural models of default face the challenge that they significantly underestimate credit spreads for corporate debt; this is the credit spread puzzle (see, e.g., Elton, Gruber, Agrawal, and Mann, 2001; Huang and Huang, 2003; Chen, Collin-Dufresne, and Goldstein, 2009). A related empirical regularity is that the puzzle is particularly strong for growth firms. Davydenko and Strebulaev (2007) show that, after controlling for standard credit risk factors, proxies of growth options are all positively and significantly related to credit spreads. Molina (2005) finds that firms with a higher ratio of fixed assets to total assets have lower yield spreads and higher ratings. Relatedly, firms with more growth options typically have lower leverage (Smith and Watts, 1992; Fama and French, 2002; Frank and Goyal, 2009). These cross-sectional features are not addressed by existing structural models because they consider firms with only invested assets.

Our model matches these facts. We show that allowing firms to be heterogeneous with respect to the importance of growth options in the values of their assets explains the aggregate credit spread puzzle, not only qualitatively, but also quantitatively. This is achieved while fitting historically reported asset volatilities and default rates for realistic debt maturities. Moreover, heterogeneity in the composition of assets can help explain cross-sectional variation of credit spreads and leverage. Our model is also consistent with observed default clustering, aggregate investment spikes and busts, and recovery rates. Additionally, we derive cross-sectional predictions for these features.

For our analysis, we develop a structural-equilibrium framework in the spirit of Bhamra, Kuehn, and Strebulaev (2010a). Thus, we embed a pure structural model of financial decisions into a consumption-based asset pricing model with a representative agent. Our model simultaneously incorporates both intertemporal macroeconomic risk (building on work by Hackbarth, Miao, and Morellec, 2006; Bhamra, Kuehn, and Strebulaev, 2010b; Chen, 2010), which has been shown to be important for explaining credit spreads and leverage, as well as expansion options. Macroeconomic shocks to the growth rate and volatility of earnings and to the growth rate and volatility of consumption arise due to switches between two states of the economy, boom and recession. The changes in the state of the economy are modeled via a Markov chain, a standard tool to model regime switches. The representative agent has the continuous time equivalent of Epstein-Zin-Weil preferences (Epstein and Zin, 1989; Weil, 1990). Therefore, how he prices claims depends on both his risk aversion and his elasticity of intertemporal substitution. Via the market price of consumption determined by the agent's preferences, we are able to link unobservable risk-neutral probabilities used in the structural model to historical probabilities. With this model, we can, therefore, study endogenously the effect of macroeconomic risk on credit spreads and optimal financing decisions.

We allow firms to have expansion options. These options are converted into invested assets when the underlying earnings process exceeds the investment boundary. We pinpoint the isolated effect of a firm's asset composition on credit risk and leverage by assuming, in the main analysis, that the exercise price of the growth option is financed through the sale of some assets in place, i.e., without additional funds being injected into the company. We also study equity-financing later in the paper. Default occurs when earnings are below the default threshold in a given regime. Shareholders maximize the value of equity by simultaneously choosing the optimal default and expansion option exercise policies. The capital structure is determined by trading off tax benefits of debt against default costs to maximize the ex-ante value of equity, i.e., the value of the firm.

The first result the model yields is that, like in other macroeconomic models, default boundaries are countercyclical, i.e., shareholders default earlier in recession than in boom. Thus, default is more likely in recession which, together with counter-cyclical marginal utilities and default costs, raises the costs of debt for all firms compared to a benchmark model without business cycle risk.

The central new feature of our model is that the asset composition alone matters significantly for the costs of debt. Two forces lead to the cross-sectional prediction that debt is particularly costly for firms with a high portion of expansion options in their assets' values. First, because options represent levered claims, firms with valuable growth options are more sensitive to the underlying earnings process than firms which consist of only invested assets. The volatility of the underlying earnings process would, consequently, underestimate the true default risk of growth firms. While the literature discusses this basic idea within equity-financed firms (Berk, Green, and Naik, 1999; Carlson, Fisher, and Giammarino, 2006), little is known about its impact on debt prices. Our structural model allows us to *jointly* analyze a firm's expansion policy and financial leverage. We show that the combination of these factors is critical for a full exploration of the quantitative implications of the riskiness of growth options on credit spreads.

The second driving force is that option values are more sensitive to macroeconomic regime changes than are assets in place. This higher sensitivity is, to some extent, another consequence of the idea that options represent levered claims. Importantly, an additional effect derives from the fact that the optimal exercise boundary of growth options increases in recession and decreases in boom. Intuitively, it is optimal to defer the exercise of an expansion option when the economy switches to recession, i.e., to wait for better times. Because the moneyness of growth options is regime-dependent, and because they represent levered claims, the continuation value of expansion options is more exposed to the macroeconomic state than the one of invested assets. Moreover, the changing moneyness causes expansion options to be less sensitive to the underlying development of the earnings process in recession than in boom, which reduces the value of the shareholders' option to defer default during bad times. Together, these effects amplify the counter-cyclicality of default thresholds for firms with a high portion of growth options. As marginal utility is high during bad times, the higher tendency to default in recession causes larger credit spreads under risk-neutral pricing for firms with expansion options than for those with only invested assets.

We then investigate the quantitative performance of the model in explaining empirically observed data. The literature suggests that an average BBB-rated firm has a 10 year credit spread in the range of $74 - 95$ basis points (bps). (This range is obtained by starting from the average bond yields reported in Davydenko and Strebulaev (2007) and Duffee (1998), and taking into account that around 35% of bond yields are due to non-default components). With our main set of parameters, a model without business cycle risk produces a mere 29 bps spread for an average firm. A standard macroeconomic model with optimal default thresholds in the spirit of Bhamra, Kuehn, and Strebulaev (2010b) or Chen (2010) implies a spread of 56 bps for average firms at issue which consist of only invested assets. Our estimate for the average BBB-rated US firm's asset composition is that total firm value is about 60% higher than the value of invested assets, which corresponds (approximately) to a Tobin's Q of 1.6.[1] For such a firm, we obtain a credit spread of about 66 bps when using optimal default thresholds, optimal expansion boundaries, and an earnings volatility such that the average asset volatility matches the one observed for BBB-rated firms. This spread is remarkably higher than the 39 bps our model implies for a firm with only invested assets. Note that the large difference arises even though leverage is kept constant; we only vary the characteristics of the assets themselves.

Empirical studies focus on aggregate data over cross-sections of firms, rather than on individual firm level data. Exploring this insight, Strebulaev (2007) is the first to show that it is crucial to consider the cross-sectional distribution of firms when relating the implications of capital structure models to empirical studies. Bhamra, Kuehn, and Strebulaev (2010b) extend this idea and investigate how the time-evolution of the cross-sectional distribution affects credit spreads and default probabilities. Following their approach, we characterize the aggregate dynamics by simulating over time a cross-section of average leverage ratios and asset composition ratios of individual firms which is structurally similar to the empirical distribution of BBB-rated firms. The average 10 and 20 year credit spreads of 81 and 103 basis points, respectively, from simulating this "true" cross-section in our model reflect their target credit spreads quite well. To solve the aggregate credit spread *puzzle*, a model needs to explain observed costs of debt while still matching historical default losses (given by the historical default probabilities and recovery rates), and asset volatilities. We consequently proceed by showing that the model-implied default rates and asset volatilities of BBB-rated firms

[1] Market values can be higher than book values also because of off-balance sheet assets, so there is, of course, a range for the asset composition of the "typical" firm.

are similar to the ones historically reported for realistic debt maturities. Hence, besides generating cross-sectional predictions for credit spreads, our model is also able to explain the credit spread puzzle.

The nature of assets, thus, has a powerful impact on costs of debt. Not surprisingly, it also affects the observed features of leverage. At initiation, we find that a firm with an average growth option optimally holds about $4-5\%$ lower leverage than one with only invested assets. Additionally, we obtain pro-cyclical optimal leverage decisions of firms, in line with Covas and Den Haan (2006) and Korteweg (2011). The reason is that the default risk is higher in recession than in boom. The negative relationship between growth options and leverage also maintains when simulating over time our model-implied cross-section of BBB-rated firms. In this simulation, however, firms deviate from their initially optimal leverage in a way such that the aggregate market leverage of the whole cross-section becomes counter-cyclical, consistent with Korajczyk and Levy (2003) and Bhamra, Kuehn, and Strebulaev (2010b).

We derive additional testable predictions when studying the aggregate dynamics of our model economy. Credit spreads and default rates are counter-cyclical, as reported in the literature. Next, aggregate investment patterns are strongly pro-cyclical, with investment spikes often occurring when the regime switches from recession to boom, reflecting the findings in the empirical investment literature (Barro, 1990; Cooper, Haltiwanger, and Power, 1999). Finally, our model makes specific cross-sectional predictions. For example, realized recovery rates are lower for growth firms.

Our paper contributes to several streams of previous research. First, the fact that growth options are empirically strongly associated with observed leverage has, of course, also prompted other explanations. The most prominent of these additional explanations, agency, comes in two primary forms: a shareholder-bondholder conflict and a manager-shareholder conflict. Appealing to the former, Smith and Watts (1992) and Rajan and Zingales (1995) suggest that debt costs associated with shareholder-bondholder conflicts typically increase with the number of growth options available to the firm due to underinvestment (Myers, 1977) and overinvestment by way of asset substitution (Jensen, 1986; see also Sundaresan and Wang, 2007).[2] According to Leland (1998), however, optimal leverage even increases when firms can engage in asset substitution. Similarly, Parrino and Weisbach (1999) conclude that stockholder-bondholder conflicts are too limited to explain the cross-sectional variation in capital structure. Childs, Mauer, and Ott (2005) show how short-term debt reduces agency costs. Hackbarth and Mauer (2010) demonstrate that the joint choice of debt priority structure and capital structure can virtually eliminate the suboptimal investment incentives of equityholders. Neither of the papers incorporates macroeconomic risk.

As for manager-shareholder conflicts, Morellec (2004) shows that agency costs of free cash flow can explain the low debt levels observed in practice, and the negative relationship between debt levels and the number of growth options; see also Barclay, Smith, and Morellec (2006). Morellec,

[2] See Lyandres and Zhdanov (2010) for an explanation for accelerated investment that does not rely on agency.

Nikolov, and Schürhoff (2009) conclude that even small costs of control challenges are sufficient to explain the low-leverage puzzle. It is still a matter of debate to what extent conflicts of interest between managers and stockholders cause the empirically observed patterns. Graham (2000), for example, tests a wide set of managerial entrenchment variables and finds only "weak evidence that managerial entrenchment permits debt conservatism" (p. 1931). In any case, our model is not inconsistent with either of these views. It offers a quantitatively important reason for the cross-sectional variation in leverage and credit spreads that derives solely from the nature of assets of firms.[3]

Second, at the core of our model is the idea from recent literature that macroeconomic (business cycle) risk matters in powerful ways for the costs of corporate debt and financial decisions, because firms are more likely to default when doing so is costly (see, e.g., Demchuk and Gibson, 2006; Almeida and Philippon, 2007; Bhamra, Kuehn, and Strebulaev, 2010b; Chen, 2010). What we add to this literature is the idea that the impact of business cycle risk depends on the asset base of a firm.

In contemporaneous and independent work, Chen and Manso (2010) set up a model similar to ours with expansion options. Their focus, however, is on the debt overhang problem, and not on explaining the credit spread puzzle or developing cross-sectional predictions – the central tasks of this paper.

Finally, our structural-equilibrium framework draws on insights from consumption-based asset pricing models (Lucas, 1978; Bansal and Yaron, 2004).

The paper proceeds as follows. In Section 2, we set up our valuation framework. We solve the model in Section 3. Section 4 discusses our parameter and firm sample choices, as well as the optimal default and expansion policies. Section 5 outlines qualitative properties of our model for the aggregate economy. We turn to the quantitative implications for BBB-rated firms in Section 6. Section 7 concludes.

2. The Model

We build a structural model with intertemporal macroeconomic risk, embedded inside a representative agent consumption-based asset pricing framework. The structural model is based on a standard continuous time model of capital structure decisions in the spirit of Mello and Parsons (1992), as extended by Hackbarth, Miao, and Morellec (2006) for business cycle fluctuations. Additionally, we explicitly model growth opportunities. Following Bhamra, Kuehn, and Strebulaev

[3] An alternative explanation for why low leverage may be optimal in the high-tech sectors is offered in Miao (2005). In his model, when a sector experiences technological growth, more competitors enter, leading to falling prices and possibly to a greater probability of default. Yet other explanations appeal to the fact that firms have the option to issue additional debt (Collin-Dufresne and Goldstein, 2001).

(2010b) or Chen (2010), embedding the model of capital structure into a consumption-based asset pricing model allows the valuation of corporate securities using the risk-neutral measure implied by the preferences of the representative agent.

The economy consists of N infinitely-lived firms with assets in place and possibly growth options, a large number of identical infinitely-lived households, and a government serving as a tax authority. We assume that there are two different macroeconomic states, namely boom (B) and recession (R). Formally, we define a time-homogeneous Markov chain $I_{t\geq 0}$ with state space $\{B, R\}$ and generator $Q := \begin{bmatrix} -\lambda_B & \lambda_B \\ \lambda_R & -\lambda_R \end{bmatrix}$, where $\lambda_i \in (0,1)$ denotes the rate of leaving state i. In the main analysis, we consider $\lambda_B < \lambda_R$, as in Hackbarth, Miao, and Morellec (2006).

The following properties hold: First, the probability that the chain stays in state i longer than some time $t \geq 0$ is given by $e^{-\lambda_i t}$. Second, the probability that the regime shifts from i to j during an infinitesimal time interval Δt is given by $\lambda_i \Delta t$. Third, the expected duration of regime i is $\frac{1}{\lambda_i}$, and the expected fraction of time spent in that regime is $\frac{\lambda_j}{\lambda_i + \lambda_j}$.

Aggregate output C_t follows a regime-switching geometric Brownian motion:

$$\frac{\mathrm{d}C_t}{C_t} = \theta_i \mathrm{d}t + \sigma_i^C \mathrm{d}W_t^C, \ i = B, R, \tag{1}$$

where W_t^C is a Brownian motion independent of the Markov chain, and θ_i, σ_i^C are the regime-dependent growth-rates and volatilities of the aggregate output. In equilibrium, aggregate consumption equals aggregate output. Hence, the above specification gives rise to uncertainty about the future moments of consumption growth.

To incorporate the impact of the intertemporal distribution of consumption risk on the representative household's utility, we assume the continuous-time analog of Epstein-Zin-Weil preferences (Epstein and Zin, 1989; Weil, 1990), which are of stochastic differential utility type (Duffie and Epstein, 1992a,b). Specifically, the utility index U_t over a consumption process C_s solves

$$U_t = \mathbb{E}^{\mathbb{P}} \left[\int_t^{\infty} \frac{\rho}{1-\delta} \frac{C_s^{1-\delta} - ((1-\gamma)\,U_s)^{\frac{1-\delta}{1-\gamma}}}{((1-\gamma)\,U_s)^{\frac{1-\delta}{1-\gamma}} - 1} \mathrm{d}s \,|\mathcal{F}_t \right], \tag{2}$$

where ρ is the rate of time preference, γ determines the coefficient of relative risk aversion for a timeless gamble, and $\Psi := \frac{1}{\delta}$ is the elasticity of intertemporal substitution for deterministic consumption paths.

The stochastic discount factor m_t then follows the dynamics

$$\frac{\mathrm{d}m_t}{m_t} = -r_i \mathrm{d}t - \eta_i \mathrm{d}W_t^C + (e^{\kappa_i} - 1)\,\mathrm{d}M_t, \tag{3}$$

with M_t being the compensated process associated with the Markov chain, and

$$r_i = -\rho\frac{(1-\gamma)}{1-\delta}\left(\frac{\delta-\gamma}{1-\gamma}h_i^{\delta-1}-1\right)+\gamma\theta_i-\frac{1}{2}\gamma(1+\gamma)\left(\sigma_i^C\right)^2-\lambda_i\left(e^{\kappa_i}-1\right) \tag{4}$$

$$\eta_i = \gamma\sigma_i^C \tag{5}$$

$$\kappa_i = (\delta-\gamma)\log\left(\frac{h_j}{h_i}\right), \tag{6}$$

see Chen (2010). h_B, h_R solve a non-linear system of equations given in the Appendix A.1, equation (A-5). r_i are the regime-dependent real risk-free interest rates, η_i the risk prices for systematic Brownian shocks affecting aggregate output, and κ_i is the relative jump size of the discount factor when the Markov chain leaves state i (and, consequently, $\kappa_j = \frac{1}{\kappa_i}$).

Credit spreads are based on nominal yields and taxes are collected on nominal earnings. To link nominal to real values such as the real interest rate introduced in the previous section, we specify a stochastic price index as

$$\frac{\mathrm{d}P_t}{P_t} = \pi\mathrm{d}t+\sigma^{P,C}\mathrm{d}W_t^C+\sigma^{P,id}\mathrm{d}W_t^P, \tag{7}$$

with W_t^P being a Brownian motion describing the idiosyncratic price index shock, independent of the consumption shock Brownian W_t^C and the Markov chain. π denotes the expected inflation rate, and $\sigma^{P,C} < 0, \sigma^{P,id} > 0$ are the volatilities of the stochastic price index associated with the consumption shock and the idiosyncratic price index shock, respectively. The nominal interest rate r_i^n is then given by

$$r_i^n = r_i+\pi-\sigma_P^2-\sigma^{P,C}\eta_i, \tag{8}$$

with $\sigma_P := \sqrt{\left(\sigma^{P,C}\right)^2+\left(\sigma^{P,id}\right)^2}$ being the total volatility of the stochastic price index.

At any point in time, the nominal earnings process of a firm follows

$$\frac{\mathrm{d}X_t}{X_t} = \mu_i\mathrm{d}t+\sigma_i^{X,C}\mathrm{d}W_t^C+\sigma^{P,id}\mathrm{d}W_t^P+\sigma^{X,id}\mathrm{d}W_t^X, \quad i=B,R, \tag{9}$$

where W_t^X is a Brownian motion describing an idiosyncratic shock, independent of the aggregate output shock W_t^C, the consumption price index shock W_t^P, and the Markov chain. μ_i are the regime-dependent drifts, $\sigma_i^{X,C} > 0$ the firm-specific regime-dependent volatilities associated with the aggregate output process, and $\sigma^{X,id} > 0$ the firm-specific volatility associated with the idiosyncratic Brownian shock. As suggested by the literature, we posit that $\sigma_B^{X,C} < \sigma_R^{X,C}$ (Ang and Bekaert, 2004).

Denote the risk-neutral measure by $\mathbb{Q}$. Following Chen (2010), the unlevered asset value is

$$V_t = X_t y_i \quad \text{for } I_t = i, \tag{10}$$

with y_i being the price-earnings ratio in state i:

$$y_i = \frac{r_j^n - \tilde{\mu}_j + \tilde{\lambda}_j + \tilde{\lambda}_i}{\left(r_i^n - \tilde{\mu}_i + \tilde{\lambda}_i\right)\left(r_j^n - \tilde{\mu}_j + \tilde{\lambda}_j\right) - \tilde{\lambda}_i \tilde{\lambda}_j}. \tag{11}$$

In Equation (11), $\tilde{\mu}_i$ are the expected growth rates of the firm's nominal earnings under the risk-neutral measure $\mathbb{Q}$

$$\tilde{\mu}_i := \mu_i - \sigma_i^{X,C}\left(\eta_i + \sigma^{P,C}\right) - \left(\sigma^{P,id}\right)^2, \tag{12}$$

and $\tilde{\lambda}_i$ are the risk-neutral transition intensities given by

$$\tilde{\lambda}_i = e^{\kappa_i}\lambda_i. \tag{13}$$

As in Bhamra, Kuehn, and Strebulaev (2010b), the price-earnings ratio in the main analysis is higher in boom, $y_B > y_R$. Finally, note that the total volatility of the earnings process is

$$\tilde{\sigma}_i = \sqrt{\left(\sigma_i^{X,C}\right)^2 + (\sigma^{P,id})^2 + (\sigma^{X,id})^2}. \tag{14}$$

The expansion option of the firm is modeled as an American call option on the earnings. Specifically, at any time $\bar{t}$, the firm can pay exercise costs K to achieve additional future earnings of sX_t for all $t \geq \bar{t}$ for some factor $s > 0$. We assume that if a firm exercises its expansion option, the option is converted into assets in place, such that the firm consists of only invested assets. The exercise of the growth option is assumed to be irreversible. At default, bondholders recover not only a fraction of the assets in place, but also a fraction of the option's value. Thus, the option can be exercised independently of the considered firm.

For the financing of investment, we present two variants. In the main analysis, we wish to abstract away from the effect of fund injections by debt- or equityholders to pay the exercise price, and instead to isolate the effect of growth options in the value of firms' assets on corporate securities. Therefore, we first assume that, at exercise, the firm pays the exercise costs K of the option by selling a part of its assets in place.[4] In detail, while exercising the option at time $\bar{t}$ entitles the firm to total future earnings of $(s+1)X_t$ for all $t \geq \bar{t}$, financing the exercise costs requires to sell a fraction $\frac{K}{X_{\bar{t}} y_{\bar{i}}}$ of these earnings, where $\bar{i}$ is the realized state of the economy at the time of exercise. Hence, the total earnings of the firm at any point in time after exercise correspond to $\left((s+1) - \frac{K}{X_{\bar{t}} y_{\bar{i}}}\right) X_t$. Second, we also consider equity financing of the exercise costs K.

The critical measure to capture the relative importance of a firm's expansion option in the value of its assets is the *asset composition ratio*. We define it as the value of the firm, divided by the

[4]Indirect financing by selling assets often occurs, e.g., when acquirers divest part of target companies' assets following takeovers (Bhagat, Shleifer, and Vishny, 1990; Kaplan and Weisbach, 1992). Of course, the model simplifies in that in reality, firms have different types of assets.

value of invested assets. Certainly, the value of the firm does not only contain the value of the invested assets and the expansion option, but also the value of the tax shield and bankruptcy costs. Nevertheless, we use this measure because the impact of the tax shield and bankruptcy costs on the ratio is relatively small, and, importantly, the direct empirical analog of the asset composition ratio is Tobin's Q.

Corporate taxes are paid at a constant rate τ, and full offsets of corporate losses are allowed. In our framework, firms are leveraged because debt allows it to shield part of its income from taxation. Once debt has been issued, a firm pays a total coupon c at each moment in time. Following the standard in the literature, we assume that firms finance coupons by injecting funds. At any point in time, shareholders have the option to default on their debt obligations, as well as the possibility to exercise an expansion option. Default is triggered when shareholders are no longer willing to inject additional equity capital to meet net debt service requirements (Leland, 1998). If default occurs, the firm is immediately liquidated and bondholders receive the unlevered asset value less default costs, reflecting the 'absolute priority' of debt claims. The default costs in regime i are assumed to be a fraction $1 - \alpha_i$ of the unlevered asset value at default, with $\alpha_i \in (0, 1]$. We suppose that recovery rates are lower in recession, i.e., $\alpha_R < \alpha_B$ (Frye, 2000). The incentive to issue debt is limited due to the possibility of costly financial distress.

Equityholders face the following decisions: First, once debt has been issued, they select the default and expansion policies that maximize equity value. Hence, both expansion and default are chosen endogenously. Second, they determine the optimal capital structure by choosing the coupon level which maximizes the value of the firm. The model does not allow restructuring of debt neither when the option is exercised nor at endogenous restructuring points. The main reason is that expansion opportunities preclude a scaling feature of the model solution.[5]

The main text presents the model and its solution for infinite debt maturity. We also solve and use the case of finite debt maturity, in which we consider the stationary environment of Leland (1998): The firm issues debt with a constant principal p and a constant total coupon c paid at each moment in time. A fraction m of the total debt is continuously rolled over. In particular, the firm continuously retires outstanding debt principal at rate mp and replaces it with new debt vintages of identical coupon and principal. Finite maturity debt is, therefore, characterized by the tuple (c, m, p). This setup leads to a time-homogenous setting. Throughout the paper, it is assumed that debt is issued at par.

[5]The scaling property states that, conditional on the current regime of the economy, the optimal coupon, the optimal default thresholds, the investment boundaries, as well as the values of debt and equity at restructuring points are all homogeneous of degree one in earnings. The absence of a scaling property impedes not only closed-form results, but also the application of numerical solution methods with backward induction. Furthermore, even formulating an ansatz for the valuation of corporate securities with expansion opportunities and restructuring requires strong assumptions on the structure of the solution, in particular on the relation between restructuring and exercise boundaries.

3. Model solution

The model is solved by backward induction. We start by calculating the value of corporate securities of a firm consisting of only invested assets, taking the capital structure, default and expansion policies as given at this point. Next, the value of the growth option, also for given capital structure and policies, is derived. We then proceed with the value of corporate securities of a firm which consists not only of assets in place, but also holds an expansion option. Finally, we obtain the expansion and default policies which simultaneously maximize the value of equity, as well as the capital structure which maximizes the firm value.

As in Hackbarth, Miao, and Morellec (2006), we assume that the optimal strategies are of regime-dependent threshold type in X (for a formal proof in the case of expansion thresholds only, see Guo and Zhang, 2004). Precisely, suppose that $\hat{D}_i$ and D_i are the default thresholds in regime $i = B, R$ of a firm with only invested assets, and of a firm with both invested assets and a growth option, respectively. X_i denotes the exercise boundary of the growth option in regime $i = B, R$. In what follows, we present the case that $D_B < D_R$, $X_B < X_R$ and $\hat{D}_B < \hat{D}_R$, i.e., the boundaries are lower in boom for both expansion and default (before and after expansion).[6] Finally, we presume that $\max\left\{D_R, \hat{D}_R\right\} < X_B$, i.e., we are interested in firms which exercise their expansion option with a positive probability, and we exclude the possibility of immediate default after expansion. The optimal default and expansion policies for relevant parameter regions satisfy the assumed ordering.

3.1. Firms with only invested assets

Let $\hat{d}_i(X), \hat{t}_i(X), \hat{b}_i(X)$ and $\hat{f}_i(X)$ denote the value of corporate debt, taxes, bankruptcy cost, and total firm value, respectively, in regime $i = B, R$. Hackbarth, Miao, and Morellec (2006) show how to solve a similar structural model.[7] The solution for our case can be found in Appendix A.2.

3.2. The value of the growth option

In order to determine the value of corporate securities of firms with both assets in place and a growth option, we now need to calculate the value of a growth option under regime switches.[8]

[6] Note that we can assume without loss of generality that $D_B < D_R$ (if not, interchange the names of the regimes). The case $D_B < D_R, \hat{D}_B < \hat{D}_R$, and $X_B > X_R$, (i.e., the exercise boundary in recession is lower than the one in boom) can be solved by analogous techniques. A changing order of the default boundaries after exercising the option, i.e., the case that $D_B < D_R$ and $\hat{D}_B > \hat{D}_R$, is not considered. Finally, the case of the presence of only one regime is presented in Appendix A.3, Case 2.

[7] Even though they do not consider regime-dependence of volatility, the basic approach remains unchanged.

[8] We emphasize that the value of the option in the ultimate solution of the model indeed depends on the default policy of the firm. Equityholders choose default and expansion policies simultaneously. The resulting interdependence between the two policies affects the value of the growth option. This effect is not explicit in the presented calculations due to the backward solution method.

Denote the value functions of the growth option in regime B and R by $G_B(X)$ and $G_R(X)$, respectively. For each regime i, the option is exercised immediately whenever $X \geq X_i$ (option exercise region); otherwise it is optimal to wait (option continuation region). This structure results in the following system of ODEs for the value function:

For $0 \leq X < X_B$:

$$\begin{cases} r_B^n G_B(X) & = & \tilde{\mu}_B X G_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 G_B''(X) + \tilde{\lambda}_B \left(G_R(X) - G_B(X)\right) \\ r_R^n G_R(X) & = & \tilde{\mu}_R X G_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 G_R''(X) + \tilde{\lambda}_R \left(G_B(X) - G_R(X)\right) \end{cases} \tag{15}$$

For $X_B \leq X < X_R$:

$$\begin{cases} G_B(X) & = & sXy_B - K \\ r_R^n G_R(X) & = & \tilde{\mu}_R X G_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 G_R''(X) + \tilde{\lambda}_R \left(sXy_B - K - G_R(X)\right) \end{cases} \tag{16}$$

For $X \geq X_R$:

$$\begin{cases} G_B(X) & = & sXy_B - K \\ G_R(X) & = & sXy_R - K \end{cases} \tag{17}$$

Whenever the process X is in the option continuation region, which corresponds to system (15) and the second equation of (16), the required rate of return r_i^n (left-hand side) must be equal to the realized rate of return (right-hand side). The latter is obtained by Ito's lemma for regime switches (see, e.g., Yin, Song, and Zhang, 2004). Here, the last term accounts for a possible jump in the value of the growth option due to a regime switch. It is calculated as the probability of a regime shift, $\tilde{\lambda}_B$ or $\tilde{\lambda}_R$, times the associated change in the value of the option. The first equation of (16) and the system (17) state the payoff of the option at exercise, since the process is in the option exercise region in these cases.

The boundary conditions are:

$$\lim_{X \searrow 0} G_i(X) = 0, \quad i = B, R \tag{18}$$

$$\lim_{X \searrow X_B} G_R(X) = \lim_{X \nearrow X_B} G_R(X) \tag{19}$$

$$\lim_{X \searrow X_B} G_R'(X) = \lim_{X \nearrow X_B} G_R'(X) \tag{20}$$

$$\lim_{X \nearrow X_R} G_R(X) = sX_R y_R - K \tag{21}$$

$$\lim_{X \nearrow X_B} G_B(X) = sX_B y_B - K \tag{22}$$

Condition (18) ensures that the option value goes to zero as the earnings approach zero. Conditions (19) and (20) represent the value-matching and smooth-pasting conditions of the value function in recession at the exercise boundary in boom. The remaining conditions (21)-(22) are the value-matching conditions at the exercise boundaries in boom and recession, respectively. The solution of this system and its derivation are given in Appendix A.3.

We remark that similar to the occurrence of default, there are two possible ways of triggering the exercise of the expansion option: Either when the idiosyncratic shock X reaches the exercise boundary X_i in a given regime, or when the regime switches from recession to boom and X lies between X_B and X_R.

The above system of ODEs (15)-(17) subject to its boundary conditions (18)-(22) determines the value of the growth option for any given pair of exercise boundaries X_B and X_R. In the full model solution, we will need to derive option values for optimal exercise boundaries of equityholders in both levered and unlevered firms. In unlevered firms, the optimal exercise boundaries are denoted X_B^{unlev} and X_R^{unlev}, respectively. These optimal exercise boundaries solve the following additional boundary conditions:

$$\lim_{X \nearrow X_R^{unlev}} G_R'(X) = sy_R \tag{23}$$

$$\lim_{X \nearrow X_B^{unlev}} G_B'(X) = sy_B. \tag{24}$$

For ease of notation, we denote the unlevered value of the growth option by G_i^{unlev}, i.e., $G_i^{unlev}(X) = G_i\left(X \left| X_B^{unlev}, X_R^{unlev}\right.\right)$. Appendix A.3 discusses the solution.

3.3. Firms with invested assets and expansion options

Using the previous results, we finally derive the value of corporate securities of a general firm, as well as the default and expansion thresholds selected by shareholders.

In each regime, the firm faces three different regions depending on the value of X: Below the default threshold, the firm is in the default region where it defaults immediately, and debtholders receive a fraction α_i of the total asset value. The firm is in the continuation region if X is between the default threshold and the exercise boundary. Finally, the exercise region is reached if X is above the exercise boundary. After exercise, the firm consists of only invested assets, endowed with the initially determined optimal coupon level. As the default policy is an ex-post policy, the optimal default thresholds now correspond to the ones of a firm with only invested assets. That is, shareholders optimally adapt their default policy upon expansion. Debtholders anticipate this change.

3.3.1. The valuation of corporate debt

Let $d_i(X)$ denote the value of corporate debt in regime $i = B, R$. An investor holding corporate debt requires an instantaneous return equal to the nominal risk-free rate r_i^n. Again, an application of Ito's lemma with regime switches shows that debt satisfies the following system of ODEs:

For $X \leq D_B$:

$$\begin{cases} d_B(X) &= \alpha_B \left(X y_B + G_B^{unlev}(X) \right) \\ d_R(X) &= \alpha_R \left(X y_R + G_R^{unlev}(X) \right) \end{cases} \tag{25}$$

For $D_B < X \leq D_R$:

$$\begin{cases} r_B^n d_B(X) &= c + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) + \tilde{\lambda}_B \left(\alpha_R \left(X y_R + G_R^{unlev}(X) \right) - d_B(X) \right) \\ d_R(X) &= \alpha_R \left(X y_R + G_R^{unlev}(X) \right) \end{cases} \tag{26}$$

For $D_R < X < X_B$:

$$\begin{cases} r_B^n d_B(X) &= c + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) + \tilde{\lambda}_B \left(d_R(X) - d_B(X) \right) \\ r_R^n d_R(X) &= c + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) + \tilde{\lambda}_R \left(d_B(X) - d_R(X) \right) \end{cases} \tag{27}$$

For $X_B \leq X < X_R$:

$$\begin{cases} d_B(X) &= \hat{d}_B \left((s+1) X - \frac{K}{y_B} \right) \\ r_R^n d_R(X) &= c + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) + \tilde{\lambda}_R \left(\hat{d}_B \left((s+1) X - \frac{K}{y_B} \right) - d_R(X) \right) \end{cases} \tag{28}$$

And, finally, for $X \geq X_R$:

$$\begin{cases} d_B(X) &= \hat{d}_B \left((s+1) X - \frac{K}{y_B} \right) \\ d_R(X) &= \hat{d}_R \left((s+1) X - \frac{K}{y_R} \right) \end{cases} \tag{29}$$

In system (25), the firm is in the default region in both boom and recession. Here, debtholders receive $\alpha_i \left(X y_i + G_i^{unlev}(X) \right)$ at default. As the default boundary in boom is lower than the one in recession, system (26) corresponds to the firm being in the continuation region in boom, and in the default region in recession. For the continuation region in boom, the left-hand side of the first equation is the rate of return required by investors for holding corporate debt for one unit of time. The right-hand side is the realized rate of return, computed by Ito's lemma as the expected change in the value of debt plus the coupon payment c. The last term captures the possible jump in the value of debt in case of a regime switch, which triggers immediate default. Similarly, equations (27) describe the case that the firm is in the continuation region in both boom and recession. The next system, (28), deals with the case in which the firm is in the exercise region in boom, and in the continuation region in recession. After exercising the option, the firm owns total assets in place of $X y_i + s X y_i - K$, reflecting the notion that the exercise costs of the growth option are financed by selling assets. The value of debt must then be equal to the value of debt of a firm with only invested assets, i.e., $d_B(X) = \hat{d}_B((s+1) X - \frac{K}{y_B})$, which is the first equation in (28). The second equation is obtained by the same approach as in (27), where the last term captures the fact that a regime switch from recession to boom triggers immediate exercise of the expansion option in this case. Finally, equations (29) describe the case that the firm is in the exercise region in both boom and recession.

The system for finite debt maturity can be found in Appendix A.5.

The boundary conditions for debt are as follows:

$$\begin{aligned}
\lim_{X \searrow D_R} d_B(X) &= \lim_{X \nearrow D_R} d_B(X) && (30)\\
\lim_{X \searrow D_R} d'_B(X) &= \lim_{X \nearrow D_R} d'_B(X) && (31)\\
\lim_{X \searrow D_B} d_B\left(X\right) &= \alpha_B \left(D_B y_B + G_B^{unlev}(D_B)\right) && (32)\\
\lim_{X \searrow D_R} d_R\left(X\right) &= \alpha_R \left(D_R y_R + G_R^{unlev}(D_R)\right) && (33)\\
\lim_{X \searrow X_B} d_R(X) &= \lim_{X \nearrow X_B} d_R(X) && (34)\\
\lim_{X \searrow X_B} d'_R(X) &= \lim_{X \nearrow X_B} d'_R(X) && (35)\\
\lim_{X \nearrow X_B} d_B\left(X\right) &= \hat{d}_B \left(\left(s+1\right) X_B - \frac{K}{y_B}\right) && (36)\\
\lim_{X \nearrow X_R} d_R\left(X\right) &= \hat{d}_R \left(\left(s+1\right) X_R - \frac{K}{y_R}\right) && (37)
\end{aligned}$$

(30) and (31) are the value-matching and smooth-pasting conditions for the debt value in boom at the default boundary in recession. Similarly, (34) and (35) are the corresponding conditions for the debt value in recession at the option exercise boundary in boom. (32) and (33) are the value-matching conditions at the default thresholds, and (36) and (37) are the value-matching conditions at the option exercise boundaries. The default thresholds and option exercise boundaries are chosen by shareholders, and, hence, we do not have the corresponding smooth-pasting conditions for debt.

The solution of this system is given in closed form in Appendix A.4.1.

3.3.2. The valuation of tax benefits

Let $t_i(X)$ denote the value of tax benefits in regime $i = B, R$. Debt coupon payments shield income from taxation. We assume full loss carry-forwards. Therefore, the value of tax benefits corresponds to the value of debt with recovery rates equal to zero and a coupon of $c\tau$. In detail, we obtain a system of equations akin to the system (25)-(29), and boundary conditions similar to (30) - (37). (32) - (33) translate into

$$\lim_{X \searrow D_i} t_i\left(X\right) = 0, \qquad i = B, R, \tag{38}$$

reflecting the loss of tax benefits at bankruptcy. At the option exercise boundary, we have that

$$\lim_{X \nearrow X_i} t_i(X) = \hat{t}_i\left((s+1)X_i - \frac{K}{y_i}\right), \qquad i = B, R, \tag{39}$$

corresponding to the conditions (36) - (37). In words, if the option is exercised, the value of the tax shield is equal to the one of a firm with only invested assets.

3.3.3. The valuation of default costs

Let $b_i(X)$ denote the value of default (or bankruptcy) costs in regime $i = B, R$. $b_i(X)$ can be calculated as the value of a debt contract with recovery rates $1 - \alpha_B$ and $1 - \alpha_R$, respectively, and a coupon of zero, as there are no continuous earnings associated with default costs.

The value-matching boundary conditions at default, (32) - (33), then correspond to

$$\lim_{X \searrow D_i} b_i(X) = (1 - \alpha_i)\left(D_i y_i + G_i^{unlev}(D_i)\right), \qquad i = B, R, \tag{40}$$

reflecting the fact that the value of default costs at the boundary must be 1 minus the recovery on the value of assets in place and the growth option. At option exercise, the value-matching conditions are given by

$$\lim_{X \nearrow X_i} b_i(X) = \hat{b}_i\left((s+1)X_i - \frac{K}{y_i}\right), \qquad i = B, R. \tag{41}$$

The intuition is that, at the exercise boundary of the option, default costs must be equal to the ones of a firm with only invested assets.

3.3.4. Firm value

Total firm value f_i in regime $i = B, R$ is given by the value of assets in place $y_i X$, plus the value of the expansion option $G_i(X)$ and the value of tax benefits from debt $t_i(X)$, less the value of default costs $b_i(X)$, i.e.,

$$f_i(X) = y_i X + G_i(X) + t_i(X) - b_i(X). \tag{42}$$

3.3.5. The valuation of equity

The levered firm value equals the sum of debt and equity values. Hence, the equity value $e_i(X)$, $i = B, R$, can be written in a closed form expression as

$$e_i(X) = f_i(X) - d_i(X) = y_i X + G_i(X) + t_i(X) - b_i(X) - d_i(X). \tag{43}$$

3.3.6. Default and expansion policies

Managers select the default and investment policies that maximize the value of equity ex-post. Denote these policies by D_i^* and X_i^*, respectively. Formally, the default policy which maximizes the equity value is determined by postulating that the first derivative of the equity value has to be zero at the default boundary in each regime. Simultaneously, optimality of the option exercise boundaries is achieved by equating the first derivative of the equity value at the exercise boundary with the first derivative of the equity value of a firm with only invested assets after expansion, evaluated at the corresponding earnings in both regimes. These four optimality conditions are smooth-pasting conditions for equity at the respective boundaries:

$$\begin{cases} e_B'(D_B^*) &=& 0 \\ e_R'(D_R^*) &=& 0 \\ e_B'(X_B^*) &=& \hat{e}_B'(sX_B^* - \frac{K}{y_B}) \\ e_R'(X_R^*) &=& \hat{e}_R'(sX_R^* - \frac{K}{y_R}). \end{cases} \tag{44}$$

We solve this system numerically.

3.3.7. Capital structure

For each coupon level c, debtholders evaluate debt at issuance anticipating the ex-post optimal default and expansion decisions of shareholders. As debt-issue proceeds accrue to shareholders, the latter do not only care about the value of equity, but also about the value of debt. Hence, the optimal capital structure is determined ex-ante by the coupon level c^* which maximizes the value of equity and debt, i.e., the value of the firm. Denote by $f_i^*(X)$ the firm value given optimal ex-post default and expansion thresholds as determined by the system (44). The ex-ante optimal coupon of this firm solves

$$c_i^* := \text{argmax}_c f_i^*(X). \tag{45}$$

As indicated in equation (45), the optimal initial capital structure depends on the current regime.

4. Results

This section summarizes the results of our model. We first describe our parameter choice in Section 4.1, and present the firm sample in Section 4.2. Next, Section 4.3 discusses the properties of the expansion option. Section 4.4, finally, analyzes the optimal default policies of individual firms with different portions of the expansion options' value in the overall value of assets. Economy-wide results are discussed in Sections 5 and 6.

4.1. Choice of parameters

Table 1 summarizes our parameter choice. Panel A shows the firm characteristics which are selected to roughly reflect a typical BBB-rated S&P 500 firm.[9] We start with an initial value of the idiosyncratic earnings X of 10. While this value is arbitrary, neither credit spreads nor optimal leverage ratios depend on this choice. As is standard in the literature, we set the tax advantage of debt to $\tau = 0.15$ (Hackbarth, Miao, and Morellec, 2006). Bhamra, Kuehn, and Strebulaev (2010b) estimate growth rates and systematic volatilities of nominal earnings in a two regime model. Their estimates are similar to those obtained by other authors who jointly estimate consumption and dividends with a state-dependent drift and volatility (Bonomo and Garcia, 1996). Hence, we choose the same growth rates and systematic volatilities of nominal earnings. The real earnings growth rates (μ_i) and volatilities ($\sigma_i^{X,C}$) correspond to their nominal counterparts net of inflation. Note that the relation $\sigma_B^{X,C} = 0.0869 < 0.1369 = \sigma_R^{X,C}$ captures the observation in Ang and Bekaert (2004) that asset volatilities are lower in boom than in recession.

Following Acharya, Bharath, and Srinivasan (2007), we assume that recovery rates fall during recession. They report that recovery in a distressed state of the industry is lower than the recovery in a healthy state of the industry by up to 20 cents on a dollar. The reason can be financial constraints that industry peers of defaulted firms face as proposed by the fire-sales or the industry-equilibrium theory of Shleifer and Vishny (1992), or time-varying market frictions such as adverse selection. We choose recovery rates as $\alpha_B = 0.7$ and $\alpha_R = 0.5$, respectively, which matches the 20 cents on a dollar difference in Acharya, Bharath, and Srinivasan (2007), and is close to the standard of 0.6 used in the literature (Hackbarth, Miao, and Morellec, 2006; Chen, 2010). Our qualitative results are insensitive to the choice of α_i as long as $\alpha_B > \alpha_R$.

Panel B shows the parameters we use to capture growth options. We select an exercise price of $K = 310$. The choice of a relatively high K is motivated by our intention to investigate firms which do not exercise their expansion option immediately. The scale parameter s for a typical firm is calibrated such that the asset composition ratio at initiation given optimal financing equals the average Tobin's Q of 1.6 in our sample of BBB-rated firms. In particular, s is set to $s = 1.89$ for firms initiated in boom, and to $s = 2.05$ for firms initiated in recession. To analyze growth firms with a larger (smaller) portion of option values in the overall value of their assets, we will later use higher (lower) scale parameters at initiation.

Panel C, finally, lists the variables describing the underlying economy. The regime-switching intensities (λ_i), the consumption growth rates (θ_i), and the consumption growth volatilities σ_i^C are estimated in Bhamra, Kuehn, and Strebulaev (2010b). We take the same values for comparability. In the described economy, the expected duration of regime B (R) corresponds to 3.68 (2.03) years, and the average fraction of time spent in regime B (R) is 0.64 (0.36). The inflation parameters are

[9] Our qualitative results do not depend on the ratings of firms.

estimated using the price index for personal consumption expenditures from the Bureau of Economic Analysis from 1947 to 2005. We obtain an expected inflation rate (π) of 0.0342, a volatility of the price index of 0.0137, and a correlation between the price index and real non-durables plus service consumption expenditures of -0.2575. These parameters imply a systematic price index volatility of $\sigma^{P,C} = -0.0035$ and an idiosyncratic price index volatility of $\sigma^{P,id} = 0.0132$.

The annualized rate of time preference, ρ, is 0.015, the relative risk aversion, γ, is equal to 10 and the elasticity of intertemporal substitution, Ψ, is set to 1.5. This parameter choice is commonly used in the literature (Bansal and Yaron, 2004; Chen, 2010).

Our calibration implies that real interest rates are $r_B = 0.0416$ and $r_R = 0.0227$ in the baseline specification. The relative decline in the value of invested assets following a shift from boom to recession is equal to 12.61%, which is similar to the one assumed in Hackbarth, Miao, and Morellec (2006).

INSERT TABLE 1 HERE

4.2. Firm Sample

Balance sheet and ratings data are collected over the period from 1995 to 2008 from Compustat. We use data for BBB-rated firms. We calculate the quasi-market leverage of a firm as the ratio of book debt to the sum of book debt and market value of equity. Tobin's Q is defined as total assets plus the market value of equity minus the book value of equity divided by total assets.[10] We delete financial and utility firms from the sample. For each firm, we calculate the average of the leverage ratios and Tobin's Qs over the observation period. Next, we cut extreme values of both average leverage and Tobin's Q at 1% to avoid that our results are driven by outliers. Our sample then consists of 717 distinct firms. Figure 1 plots the resulting data points. For the entire cross-section of BBB-rated firms, the mean leverage is 41.83%, and the mean Tobin's Q (asset composition ratio) is 1.59.

INSERT FIGURE 1 HERE

[10] In these definitions, we follow, e.g., Baker and Wurgler (2002), Fama and French (2002) and Daines, Gow, and Larcker (2010). Book debt is total assets (item 6, AT) minus book equity. Book equity is total assets minus total liabilities (item 181, LT) minus preferred stock (item 10, $PSTKL$, replaced by item 56 when missing, $PSTKRV$) plus deferred taxes (item 35, $TXDITC$) plus convertible debt (item 79, $DCVT$). The market value of equity is given by the closing price (item 24, $PRCC\_F$) times the number of common shares outstanding (item 25, $CSHO$).

4.3. Properties of the expansion option

To understand the implications of our model for credit spreads, it is instructive to first consider some properties of the expansion option.

Figure 2 depicts the equity value maximizing exercise policy of the expansion option in a typical firm initiated in boom. Recall that the expansion policy is simultaneously determined with the default policy.

INSERT FIGURE 2 HERE

The area above the dashed line is the exercise region in recession, and the area below the dashed line represents the continuation region. In boom, the regions are defined analogously with respect to the solid line. The graph is drawn for optimal leverage. Exercising the option at time $\bar{t}$ entitles the firm to total future earnings of $(s+1)X_t$ for all $t \geq \bar{t}$. As expected, the endogenous exercise boundaries decrease with s. For example, consider initiation in boom: With a scale parameter of $s = 1.89$ (which induces an asset composition ratio of 1.6 at initiation), the corresponding optimal option exercise boundaries are $X_B^* = 18.14$ and $X_R^* = 19.14$. Setting s to 2.73 creates a growth firm with an asset composition ratio of 2.2, and optimal option exercise boundaries of $X_B^* = 12.87$ and $X_R^* = 13.55$, respectively. Importantly, Figure 2 also shows that the expansion option is exercised at lower levels of the idiosyncratic earnings X in boom than in recession. Intuitively, the main reason is that the value of the real option of waiting is higher in recession due to the potential switch to boom with a higher valuation of earnings.[11] The same qualitative option value properties also hold at non-optimal leverage levels.

Figure 3 plots the value of the expansion option as a function of the earnings X, using jointly optimal expansion and default policies.

INSERT FIGURE 3 HERE

Obviously, the option's value is affected by the current regime. When the asset value jumps due to a regime switch, so does the value of the option. Critically, relative value changes of expansion options are higher than relative value changes of assets in place when the regime switches. The reasons are that options represent levered claims, and that the endogenous exercise boundary is higher in recession than in boom, as shown in Figure 2.[12]

[11] The regime dependent volatilities and default thresholds also affect optimal exercise boundaries in boom and recession. We find that the valuation of earnings is the dominating effect for reasonable parameter values.

[12] Relative value changes are determined in Appendix A.3. In untabulated results, we confirm numerically that the relative value changes are indeed higher for expansion options than for the underlying assets in place for plausible parameter values.

Additionally, Figure 3 shows that both option value functions are convex, but the value function in boom is steeper than the one in recession. Therefore, the expansion option's value is less sensitive to the underlying earnings in recession than in boom. Intuitively, the exercise boundary increases and the earnings' drift decreases in recession which drives options out-of-the money. As a consequence, an expansion option represents a less levered claim in bad times. While in recession the volatility of X is higher, the sensitivity of a growth option's value to changes in the earnings is lower. As discussed in the next section, this lower sensitivity attenuates the increase in the equityholder's default option due to a higher volatility of X during recession.

4.4. Optimal Default Policy

This section explains how the optimal default policy is affected by the presence of growth options in the value of firms' assets. In order to keep the intuition tractable, we do not comment on the (minor) impact of the exercise boundaries on default thresholds, which arises due to the simultaneous optimization of the expansion and default policy.

For all firms – those with and those without an expansion option – the optimal default policy is determined by recognizing that, at any point, shareholders can either make coupon payments and retain their claim together with the option to default, or forfeit the firm in exchange for the waiver of debt obligations. When the economy shifts from boom to recession, the present value of future earnings declines mainly because firm earnings have a lower drift, and because they become both more volatile and more correlated with the market. This present value decline reduces the continuation value (the expected value from keeping the firm alive) for equityholders, inducing them to default earlier (at higher earnings levels) in recession. We will refer to this effect as the *value effect*. On the other hand, a high earnings volatility in recession makes the option to default more valuable, which defers default in bad times. This is the *volatility effect*. As in the models for invested assets of Bhamra, Kuehn, and Strebulaev (2010b) and Chen (2010), the value effect usually dominates the volatility effect, generating higher default thresholds in recession, i.e., leading to counter-cyclical default thresholds. Counter-cyclical default thresholds together with a high volatility in bad times imply counter-cyclical default probabilities, consistent with empirical evidence (Chava and Jarrow, 2004; Vassalou and Xing, 2004). Additionally, default losses are empirically reported to be higher in recession because many firms experience poor performances during such times. Combined with higher marginal utilities in bad times, these mechanisms raise the present value of expected default losses for bondholders which leads to higher credit spreads and lower optimal leverage ratios than in standard contingent claim models.

Figure 4 draws the equity value maximizing default policy of levered firms initiated in boom. The graph shows default thresholds for a range of asset composition ratios. Leverage is held constant at 41.83%.[13] The solid line represents the default threshold in boom, and the dashed line the one in recession. For a firm with only invested assets the optimal default thresholds correspond to $D_B^* = 2.25$ and $D_R^* = 2.46$, for an average firm with an asset composition ratio of 1.6 to $D_B^* = 2.81$ and $D_R^* = 3.19$, and for a growth firm with an asset composition ratio of 2.2 to $D_B^* = 2.98$ and $D_R^* = 3.41$. In the no-default region above the line corresponding to a given regime, the continuation value for equityholders exceeds the default value and it is optimal for shareholders to keep injecting funds into the firm.

INSERT FIGURE 4 HERE

Two points from Figure 4 are particularly noteworthy. First, the optimal default thresholds increase as the asset composition ratio increases, inducing a higher default probability. This finding evolves from the observation that growth options represent levered claims which are relatively more sensitive than invested assets to a given decrease in X. Second, while all firms are more likely to default in recession than in boom, the increasing distance between D_B^* and D_R^* for larger asset composition ratios indicates that the counter-cyclicality of default boundaries is particularly pronounced for growth firms. The reason is that due to the higher relative value change of growth options upon a regime switch, the value effect is stronger for a firm with a high asset composition ratio. Additionally, because options represent less levered claims in recession than in boom, the increase in the equityholders' default option - due to the higher volatility of X when the regime switches to recession - is attenuated for growth firms. In other words, the volatility effect, which tends to decrease the distance between the default thresholds, is weaker for firms with larger expansion options.

5. Aggregate dynamics of leverage, asset composition, investment and defaults

In order to validate our structural-equilibrium framework with intertemporal macroeconomic risk and investment, we analyze the dynamic properties of our model-implied economy. In this section, we qualitatively compare the aggregate predictions for the entire economy to empirically reported capital structure, investment, and default patterns. In Section 6, we quantitatively explain observed credit spreads and leverage ratios of the subset of BBB-firms.

[13] When the scale parameter is changed but the coupon is left constant, default thresholds are not directly comparable. The reason is that the total asset value increases with s for every X. Taking constant leverage assures that the considered coupon changes consistently with the increase in the total asset value when we alter s.

5.1. Simulation

We generate a dynamic economy of average firms implied by our model. We consider 1,000 identical firms with infinite debt maturity. Initially, each firm's earnings are $X = 10$, and the option scale parameter is assumed to be $s = 1.89$ if the firm's initial regime is boom, and $s = 2.05$ otherwise. These choices of s imply an asset composition ratio of 1.6 in both states at initiation, given optimal leverage. Firms receive the same macroeconomic and inflation shocks, but experience different idiosyncratic shocks. Each firm observes its current earnings as well as the current regime on a monthly basis and behaves optimally: If the current earnings are below the corresponding regime-dependent default threshold, the firm defaults immediately; if the current earnings are above the corresponding regime-depending option exercise boundary, the firm exercises its expansion option; otherwise, the firm takes no action.

In our model, firms have a growth option which can only be exercised once. To maintain a balanced sample of firms, and to avoid that the average asset composition ratio is systematically trending towards the one of a firm with only invested assets when we simulate the economy over time, we exogenously introduce new firms. In particular, we substitute each defaulted or exercised firm by a new firm whose growth option is still intact. New firms have initial earnings of $X = 10$, and an option scale parameter s according to the current regime as described above.

To ensure convergence to the long-run steady state, we first simulate the economy for 100 years. The starting period for the reported results is the final period of the first 100 years of simulation. Next, we simulate the model for 200 years and present the aggregate dynamics.

5.2. Results

We start by discussing the cyclicality of leverage. Hackbarth, Miao, and Morellec (2006) generate counter-cyclical optimal leverage ratios in their macroeconomic model. As in our framework, the optimal coupon rate of debt initiated in boom exceeds the one in recession. At the same time, the value of assets is greater in boom. The second effect dominates the first, generating the counter-cyclicality in optimal leverage. We additionally incorporate the empirical fact that asset volatility is regime-dependent. Because the latter decreases in boom and increases in recession, our optimal coupon rate varies more than in Hackbarth, Miao, and Morellec (2006) when the regime changes. With this extension, the change in the value of optimal debt dominates the change in the value of assets, generating pro-cyclical optimal leverage ratios for realistic parameter values, in line with Covas and Den Haan (2006) and Korteweg (2011). Figure 5 plots the simulated market leverage in the economy. Shaded areas represent recessions. Even though our optimal initial leverage ratios are pro-cyclical, the simulated time series shows that actual aggregated market leverage is counter-cyclical. The reason is that when firms are stuck with the debt issued at initiation, the equity value declines more than the debt value during recessions which tends to increase leverage in bad times. This prediction conforms to Korajczyk and Levy (2003) who show that unconstrained firms' leverage ratios vary counter-cyclically.

INSERT FIGURE 5 HERE

Figure 6 shows the time series of the aggregate asset composition ratio in the simulated economy. As expansion options are more sensitive to the underlying stochastic processes than invested assets, the ratio behaves pro-cyclically, as reported in the literature.

INSERT FIGURE 6 HERE

We investigate aggregate default rates in Figure 7. Simulated default rates are counter-cyclical, consistent with the empirical fact that most defaults occur during economic recessions. Additionally, the graph shows several spikes in default rates that occur right at the time when the economy enters into a recession, consistent with the empirical evidence in Duffie, Saita, and Wang (2007) and Das, Duffie, Kapadia, and Saita (2007) (see, e.g., around years 50 and 90). Recall that defaults can occur because either the idiosyncratic earnings reach the default threshold in a given regime, or due to a change of the macroeconomic regime from boom to recession. The clustered default waves occur due to an increase in firms' default thresholds upon such a regime change. All firms between the two thresholds default simultaneously when the regime switches to recession, even though their earnings do not exhibit instantaneous regime-induced changes. After such waves of default, the default frequency tends to remain high during recessions.

As a refinement of this general result, we expect that the tendency to default during recession should be particularly pronounced for firms with high expansion options. This prediction is suggested by the fact that the degree of counter-cyclicality of default thresholds is positively related to the initial asset composition ratio. We investigate the propensity to default during recession in a dynamic, simulation-based setting by counting default rates of two separate aggregate economies. The first one is designed as above, consisting of firms with both assets in place and growth options, such that the asset composition ratio at initiation is 1.6. The second setting consists of firms with only invested assets. To construct a number of cross-sectional distributions of firms, we first simulate 20 dynamic economies for 10 years. Using each economy obtained at the end of the first 10 years, the default rates in both regimes are observed for 50 subsequent simulations of the following 20 years, resulting in a total of 1,000 simulations. The average percentage of defaults which occurs during recession is then calculated.[14] We find that in the first economy, on average,

[14]The distance to default in the aggregate economy of firms with only invested assets is trending over time. The reason is that firms which default are replaced, but there are no option exercises after which well performing firms could be replaced. Consequently, we do not compare absolute default rates of the two economies, but rather the fraction of defaults occurring in each regime.

75.41%, 76.79%, and 77.66% of total defaults of firms with assets in place and growth opportunities occur during recession over 5, 10, and 20 years, respectively. In the economy where firms only have invested assets, the corresponding numbers are considerably smaller at 66.40%, 71.66%, and 73.71%, respectively.

This finding is also related to the observation that, on average, growth firms have lower recovery rates than value firms (Cantor and Varma, 2005). The standard argument offered by Shleifer and Vishny (1992) is that growth firms as potential buyers of growth assets have little cash relative to the value of assets. Hence, they are likely to be themselves credit constrained when other growth firms sell their assets upon default, which lowers recovery rates. Our model delivers an alternative explanation: We show that growth options in the value of firms' assets create a propensity to default during recession, when recovery rates are low.

INSERT FIGURE 7 HERE

A significant literature suggests that business cycle shocks common to all firms play a crucial role in explaining aggregate investment. In particular, there is evidence that aggregate investment is characterized by both episodes of very intense investment activity and periods of very low investment activity (Doms and Dunne, 1998; Oivind and Schiantarelli, 2003). Moreover, aggregate investment and the probability of investment spikes are strongly pro-cyclical (Barro, 1990; Cooper, Haltiwanger, and Power, 1999). Our model reflects these features. First, when the regime switches from recession to boom, firms in the region between the two investment boundaries exercise their expansion option simultaneously by investing K. Figure 8 shows that investment spikes often occur upon such regime switches (see for example around year 35, or year 60). After these spikes, simulated investment rates tend to remain high during boom due to the positive drift of the earnings. Hence, we observe pro-cyclical investment spikes followed by higher investment activity during booms. At the other end, investment activity often dries out when the economy switches from boom to recession, because the optimal exercise boundary jumps up and the expected earnings' drift turns negative. Our model also predicts that observed investment waves should be mainly driven by firms with high expansion options.

INSERT FIGURE 8 HERE

Finally, we plot simulated average credit spreads in Figure 9. Credit spreads are calculated as $(c/d_i(X))-(c/RF)$, where RF is the value of a risk free bond with an identical coupon.[15] Consistent with the empirical literature (Fama and French, 1989), we find counter-cyclical credit spreads.

[15] RF is given by

$$RF = \frac{c\left(r_j^n + \tilde{\lambda}_i + \tilde{\lambda}_j\right)}{r_i^n r_j^n + r_j^n \tilde{\lambda}_i + r_i^n \tilde{\lambda}_j}. \tag{46}$$

When the economy stays in boom, credit spreads tend to decline as distances-to-default increase due to the positive expected drift of the earnings and the lower default threshold. Conversely, in recession, credit spreads rise as distances-to-default tend to decline and the volatility increases.

INSERT FIGURE 9 HERE

6. Quantitative implications and empirical predictions

In this section, we discuss the quantitative implications and empirical predictions of our model. The attention is restricted to BBB-rated firms since it has been argued that the pricing of very high-grade investment firms is dominated by factors other than credit risk such as liquidity risk or a tax component (Longstaff, Mithal, and Neis, 2005; De Jong and Joost, 2006). We start by determining target observed average credit spreads. Duffee (1998) estimates an average yield spread in the industrial sector between BBB-rated bonds and Treasury yields of 198, 148, and 149 bps for bonds with a mean maturity of 21 years (long), 8.9 years (medium), and 4.7 years (short), respectively. Davydenko and Strebulaev (2007) report somewhat lower spreads of 143 bps for bonds with $15-30$ years (long), 115 bps for $7-15$ years (medium), and 115 bps for $1-7$ years maturity, respectively.[16] From these spreads, we subtract 35.5% to reflect the results in Longstaff, Mithal, and Neis (2005) and Han and Zhou (2011) who find non-default components in BBB bond yields of 29% and 42%, respectively. We arrive at a plausible target range of around 92 to 128 bps for long maturities, 74 to 95 bps for medium maturities, and 74 to 96 bps for short maturities.[17] Panel A in Table 2 tabulates these target credit spread ranges. In Panel B, we also report empirical default rates of BBB-rated debt over 5, 10, and 20 years from Moody's (2010).

INSERT TABLE 2 HERE

We discuss the implications of our model for credit spreads and leverage along two dimensions. First, we follow the traditional way of investigating a typical individual firm. Second, we implement an approach similar to the one proposed by Bhamra, Kuehn, and Strebulaev (2010b) where credit spreads and leverage ratios are calculated as cross-sectional averages based on a simulation of the empirical distribution of BBB-rated individual firms.

[16] The estimates of short and medium maturities in Huang and Huang (2003) are higher because of the embedded call options in the corporate bond sample and the inclusion of two recessions with high spreads.

[17] We recalculate target ranges by subtracting the absolute non-default component for BBB firms of 61.8 bps reported in Han and Zhou (2011), or by subtracting the 29% reported in Longstaff, Mithal, and Neis (2005) for an earlier sample period. Our model's performance does not depend on the exact definition of targets.

6.1. Credit spreads

6.1.1. Typical firm with endogenous default boundary

Credit spreads for various models on newly issued corporate debt are calculated in Table 3 for 5 (short), 10 (medium), and 20 (long) years maturity.[18] We follow the standard approach in structural models by calibrating the idiosyncratic earnings volatility such that the total asset volatility is approximately 25% in each model, the average asset volatility of firms with outstanding rated corporate debt (Schaefer and Strebulaev, 2008). Additionally, we fix leverage at the average ratio of 41.83% in our BBB-firm sample.

Importantly, the default boundaries and expansion thresholds are assumed to be chosen optimally by equityholders, as we are interested in whether our model can generate both realistic prices of corporate claims and realistic endogenous default and expansion rates. Specifying default boundaries exogenously such that a model's actual default probabilities match the data (as done in Chen, Collin-Dufresne, and Goldstein (2009) or Huang and Huang (2003)) would not only substantially dilute the value of the option to default, but would also distort the value of the expansion option because the latter depends on the default policy.

It is well-known that structural models of default typically generate credit spreads which are too low compared to their empirical counterpart. To illustrate this point, we first analyze the model without business cycle risk in Panel A of Table 3. The expected drifts and systematic volatilities of earnings and consumption are set equal to their unconditional means. Panel A shows credit spreads for different maturities of the standard structural model of Leland (1998). The empirical target credit spreads in Table 2 are about 5 times larger for the short maturity, and about 3 times larger for the medium and long term than those predicted by the structural model.

INSERT TABLE 3 HERE

Bhamra, Kuehn, and Strebulaev (2010b) and Chen (2010) derive structural multi-regime models for typical firms which consist of only invested assets. We closely replicate their approach for an average firm within a two-regime model. To match the asset volatility of 25%, the idiosyncratic earnings volatility is set to $\sigma^{X,id} = 0.21$. Panel B reports unconditional credit spreads, calculated as a weighted average of the state-dependent credit spreads, where the weights correspond to the long-run distribution of the Markov chain. For comparability to our setting with expansion options, the results without debt restructuring are presented. While the credit spreads for typical firms of 35, 56, and 78 bps for 5, 10, and 20 years maturity, respectively, are clearly higher than in the one regime case, they are still considerably below their targets.[19]

[18] For the value of a finite maturity risk-free bond see Appendix A.5, formula (A-117).

[19] Bhamra, Kuehn, and Strebulaev (2010b) use higher recovery rates, lower leverage and do not model the impact of principal repayments on default thresholds, which results in marginally lower credit spreads in their static case.

Next, we investigate our model with expansion options for a typical BBB-rated firm. Note that for a given idiosyncratic earnings volatility, firms with different asset composition ratios have different total asset volatilities due the inherent leverage of their expansion option. Moreover, a firm's asset volatility is not constant over time, as its option's moneyness changes when X moves towards or away from the exercise boundary. To obtain the average volatility for a certain rating class, the standard approach in the literature is to average the calculated asset volatilities over all firms with the same rating (Vassalou and Xing, 2004; Duan, 1994; Schaefer and Strebulaev, 2008). We calibrate the idiosyncratic volatility $\sigma^{X,id}$ to the empirically reported average asset volatility of 0.25: Given an idiosyncratic volatility $\sigma^{X,id}$, we simulate model-implied samples of BBB-rated firms over 10 years, and calculate the resulting average asset volatility. (Details on the simulation can be found in Appendix A.6.1.) The calibration yields $\sigma^{X,id} = 0.168$ which ensures that the average asset volatility of our simulated BBB-rated firms with expansion options corresponds to its empirical counterpart.[20]

Panel C of Table 3 shows the resulting credit spreads for typical firms. Several aspects are noteworthy about these results. Our model increases the unconditional credit spreads of an average firm for 5, 10, and 20 years from 18 bps to 44 bps (+144%), from 29 bps to 66 bps (+128%), and from 41 bps to 81 bps (+98%), respectively, compared to the one regime model in Panel A. To understand this large effect, recall first that macroeconomic models generate larger credit spreads than one regime models because recessions are times of high marginal utility, so that default losses that occur during these times will affect investors more. An important economic implication is that the average duration of bad times in the risk-neutral world is longer than in the actual world. Since the representative agent uses risk-neutral and not actual probabilities to account for risk and to compute prices, credit spreads are larger and the agent behaves more conservatively than historical default losses imply. Second, if firms have a higher tendency to default in recession, this discrepancy will increase due to the higher risk premium. Our model shows that because of the strong sensitivity of option values to regime switches, and because they are less sensitive to the underlying earnings during recession, the counter-cyclicality of default thresholds is more pronounced for firms with larger growth options. The resulting stronger counter-cyclicality of the default probability of growth firms thus drives up their credit spreads. As can be seen in row 2 of Panel C, the credit spreads for an average firm, consisting of both invested assets and growth options, are 44 bps, 66 bps, and 81 bps for debt maturities of 5, 10, and 20 years, respectively.

Chen (2010) obtains larger 10 year credit spreads in a model with 9 states and a dynamic capital structure, but uses higher leverage, and a cash flow volatility which induces a much higher asset volatility than empirically observed.

[20] We also repeat this exercise with different specifications, such as alternative simulation length and debt maturity. The resulting idiosyncratic volatilities are fairly insensitive to these variations. An alternative approach is to calibrate the idiosyncratic volatility to the cumulative default probability of BBB-rated firms (Chen, 2010). This procedure, however, usually leads to asset volatilities which are higher than the ones empirically observed.

This is, respectively, 26%, 18%, and 5% higher than the credit spreads of an average firm in the standard macroeconomic model with only invested assets.[21]

Besides the fact that they generate too low credit spreads, another problem of existing structural models is that the implied term structure of credit spreads at initiation is much steeper than its empirical counterpart for a typical firm. The reason is that the implied spreads are particularly low at the short end. Most existing studies in the macroeconomic model literature use the default thresholds of infinite maturity debt (that is, debt without principal repayments) to numerically calculate the risk-neutral default probability for each maturity. As the credit risk literature identifies firms' debt maturity as an important determinant of credit risk (Gopalan, Song, and Yerramilli, 2010; He and Xiong, 2011), we endogenously derive optimal default thresholds also for less than infinite debt maturity following the approach of Leland (1998). Due to the continuous principal repayments, these thresholds are considerably higher for short maturities than for infinite debt, resulting in larger credit spreads at the short end. The resulting term structure of credit spreads for an average firm in Panel A, B, and C is consequently flatter and, hence, closer to the shape observed in target spreads than when using default thresholds of infinite maturity debt.[22]

The rows in Panel C of Table 3 identify the cross-sectional relationship between the asset composition ratio and credit risk. To tease out the effect of growth options on credit spreads, we vary the asset composition ratio by altering s. As raising s increases the value of the expansion option, we simultaneously adapt the coupon to maintain a constant leverage of 41.83%.[23] This exercise shows that the asset base of the firm is an important driver of credit risk, implying a positive relationship between the portion of growth options in the value of a firm's assets and the costs of debt. In particular, altering the asset composition ratio of a firm from 1 to 2.2 increases credit spreads by about 55% to 95%, depending on the debt maturity. This effect is remarkable given that we solely vary the assets' characteristics. It arises for two reasons in our model. First, because options are levered, and due to the endogenous investment boundary, expansion options are more sensitive to the underlying uncertainty, and, hence, more volatile. This higher volatility drives up the default probability of growth firms. Second, a higher portion of the expansion option's value in the overall asset value of a firm induces a higher counter-cyclicality of the default probability,

[21] We cannot directly compare the results for invested assets in Panel C to the ones for average firms in Panel B, even though the latter consist of only invested assets. The reason is that in our model, the idiosyncratic volatility is calibrated such that the asset volatility of the entire sample of BBB-rated firms matches 0.25, whereas in Panel B, $\sigma^{X,id}$ is chosen such that firms with only invested assets have an asset volatility of 0.25.

[22] We use default boundaries for the appropriate debt maturities in both Panels B and C in order to highlight the pure effect of expansion options on credit spreads.

[23] Alternatively, changing both s and K to alter the asset composition qualitatively retains the aggregate and cross-sectional predictions. Holding s constant while only varying K implies large decreases in the option exercise boundaries for relatively small increases of the asset composition ratio. In the extreme, a firm with a very low K will exercise its expansion option almost immediately; in essence, credit spreads then virtually mirror those of a firm with only invested assets, diluting the model's cross-sectional predictions. Note also that any variation in K changes the costs of investment. By only varying s, we instead avoid that our results are driven by different sizes of the expected financing in case of equity-financed investment costs.

which raises expected default costs. The higher default probability and larger default costs both increase the costs of debt for growth firms.

Note that while firms with growth options generally have a higher credit spread than firms with only invested assets (ceteris paribus), credit risk is concave in the asset composition ratio. This concavity occurs because firms with a larger asset composition ratio are closer to their exercise boundary, where credit spreads also reflect that the asset volatility and the counter-cyclicality of the default thresholds will decrease when a firm exercises its expansion option.

Our predictions are qualitatively consistent with empirical findings. For example, Davydenko and Strebulaev (2007) find that market-to-book asset values, the ratio of research and development expenses to total investment expenditure, and one minus the ratio of net property, plant, and equipment to total assets are all significantly and positively related to credit spreads (Table VI on p. 2652). Similarly, Molina (2005) documents that firms with a higher ratio of fixed assets to total assets have lower bond yield spreads and higher ratings (Table II on p. 1438). This evidence implies that, empirically, even after controlling for most factors relevant to credit risk in standard structural models, credit spreads are higher for growth firms. Hence, while an average firm with valuable growth options exhibits, for example, a different tax advantage of debt or payout ratio than a firm which only consists of invested assets, simple variation of such input parameters would not explain these findings. What is needed to address the aggregate puzzle and the mentioned cross-sectional evidence is a model which generates higher explained credit risk than standard models for a *given* level of input parameters. Our model delivers this result.

6.1.2. True cross-section

The previous section calculates credit spreads of a typical individual firm which is consistent with the historically observed average input parameters of firms in the same rating class of which the individual firm is representative.

In this section, inspired by the work of Bhamra, Kuehn, and Strebulaev (2010b), we employ a simulation approach to capture the dynamics of the cross-sectional distribution of leverage and asset compositions of BBB-rated firms. The central insight of this approach is that BBB-rated firms are very different with respect to their firm characteristics such as the asset composition ratio and leverage, and that credit spreads and default rates are highly non-linear in these characteristics. Moreover, the previous section considers credit spreads solely at debt issuance points, when the principal corresponds to the market value of debt. The majority of empirically reported spreads are, however, based on observations made at times when debt is not being issued. To capture the impact of these issues, it is important to calculate credit spreads and default rates for a simulated sample of firms which matches the observed empirical distribution, i.e., the true observed cross-section of BBB-rated companies. The resulting average of simulated credit spreads can then be compared to the empirical average credit spread. Simultaneously, the approach allows us to verify whether the default probabilities implied by our model correspond to the reported historical default probabilities of BBB-rated firms.

To obtain the implications of the true cross-section of BBB-rated firms, we start by generating a distribution of firms implied by the model. In particular, we set up a grid of optimally leveraged firms with scale parameters s from 0 up to the largest possible value such that the option is not exercised immediately. The step size is 0.05, and 50 identical firms are considered for each value of the option scale parameter. Earnings paths of all firms are then simulated forward over 10 years, resulting in a model-implied economy populated by more than 3000 firms. This economy has a broad range of leverage ratios and asset composition ratios.

In a second step, we match our historical distribution of BBB-rated firms with its model-implied counterpart. For each observation in the average empirical cross-section, we select the firm in our model-implied economy with the minimum distance regarding the percentage deviation from the target average market leverage and asset composition ratio. The matching is generally very accurate. Considering a debt maturity of 10 years yields an average Euclidean distance of 0.0648, with the 85%-quantile being 0.0865.[24] That is, on average, only 15% of the firms are matched with the root of the sum of the squared percentage deviations being larger than 8.65%.[25] Note that while our initial model-implied economy potentially contains firms with different ratings, the described matching procedure allows us to construct a cross-sectional distribution of model-implied firms which closely reflects its empirical BBB-rated counterpart.

Next, earnings paths of the 717 matched BBB-model-firms are simulated forward for 20 years on a monthly basis. This simulation is repeated 50 times.

The outcome of both the matching and the forward simulation of the matched sample also depends on the particular realizations of the idiosyncratic shocks and the states of the economy in the first simulation step. Hence, to explore the distributional properties of our results, the entire procedure is conducted 20 times, which results in a total of $1,000$ simulations. Details on the simulation are given in Appendix A.6.2.

Panel D of Table 3 summarizes the results. The average credit spreads, calculated during 5 years after the matching, are 60 bps for 5 years, 81 bps for 10 years, and 103 bps for 20 years.[26]

[24] Other debt maturities yield virtually identical results for the matching accuracy.

[25] The market leverage is matched with an average distance of 0.0248. The average percentage distance of the asset composition ratio of 0.0549 is larger. This number is driven by a few firms with unusually high asset composition ratios. As they would optimally exercise their expansion option immediately in our model, these firms are matched with model firms with a somewhat lower asset composition ratio. We expect a minor impact of this limitation on our results, because firms with unusually high asset composition ratios also have very low leverages, and, hence, are not driving our average credit spreads.

[26] We follow Bhamra, Kuehn, and Strebulaev (2010b) in measuring average credit spreads over a 5 year period. During longer periods, many firms could deviate substantially from the initial average distribution, and would, therefore, not be BBB-rated anymore.

Hence, our model closely matches the historical levels reported in Table 2 for 10 and 20 years. 5 year credit spreads are somewhat lower than their target. We also measure the cyclicality of credit spreads. Average 10 year credit spreads, for example, are 58 basis points during boom, and 112 during recession. As expected, they are strongly counter-cyclical.

Importantly, average credit spreads for the simulated true cross-section are considerably higher than the ones of a typical firm at initiation. There are two reasons for this result. First, some firms will be near default, and credit spreads are convex in the distance to default. Second, the market value of debt corresponds to the principal at initiation. In practice, however, firms are not at initiation most of the time. The actual market value of debt will, therefore, often underestimate the burden from the principal repayments, and especially so for firms approaching their default boundary. The reason is that the market value can hardly go beyond the principal as it is bounded above by the value of riskless debt, but can easily reach values below the principal when earnings deteriorate. Our simulation of the true cross-section captures these asymmetric deviations over time, resulting in higher average credit spreads than those of firms observed at initiation. Compared to Bhamra, Kuehn, and Strebulaev (2010b), the additional credit spreads generated from simulating the true cross-section are lower, because we do not incorporate debt restructuring.

To verify whether our model generates default rates corresponding to the empirically reported default frequencies for realistic debt maturities, we also count cumulative default rates in the simulated true cross-section. The model-implied average and median cumulative default rates over several years are reported in each Panel of Table 4. Panel A presents default rates over 5, 10, and 20 years from simulations with firms issuing infinite maturity debt. Panels B, C, and D show default rates from simulations with firms issuing finite maturity debt. Due to the principal repayments, default thresholds of firms with finite maturity debt are considerably higher than those of firms with infinite maturity debt. Note that simulated credit spreads are consistent with a range of realized ex-post default rates, as observed default rates vary depending on a particular realization of good and bad states. Therefore, we also report the 25% and 75% percentiles of the distribution.

Empirically, Datta, Iskandar-Datta, and Patel (2000) report a mean maturity of IPO bonds of 12 years, Guedes and Opler (1996) document an average maturity of 12.2 years for seasoned debt offers, and Davydenko and Strebulaev (2007) measure a mean time to maturity of BBB-bonds in the industrial sector of 9.51 years. Panel C of Table 4 shows that when assuming that firms have a debt maturity of 10 years, our model-implied median default rates over 5, 10, and 20 years are very close to the historical default probabilities observed from 1920 to 2009 reported in Table 2. Hence, for a realistic debt maturity, our median economy is consistent with historical default frequencies of BBB-rated firms. The average default rates are somewhat larger than their targets due to a few realizations with long sojourn times in recession, resulting in high default rates.[27] Panels A and

[27] The standard deviation of the sojourn times generated by Markov chains is quite large. In our model, long sojourn times in recession cause high default rates for some sample paths. As default rates are non-linear in the distance to default, long sojourn times in boom do not counterbalance these high rates.

D show that while the generated rates tend to be too low in Panel A, but too large in Panel D, historical default frequencies still fall within the 25% to 75% range of model-implied median default rates for most years.

The large difference between Panel A and D in both average and median default rates illustrates that debt maturities and the associated default thresholds have an important effect on model-implied default rates. It is, therefore, important to incorporate a realistic debt maturity when calibrating models with endogenous default thresholds.

INSERT TABLE 4 HERE

In sum, our results demonstrate that the average credit spreads implied by our model for the true cross-section are simultaneously consistent with historically observed average asset volatilities, and, especially for typical debt maturities, with default rates reported for BBB-rated firms.

6.2. Leverage

This section analyzes the features of leverage ratios resulting from our model. We first investigate how growth options affect the initial choice of optimal leverage in our model. At initiation, a firm consisting of only invested assets has an optimal leverage which is between 4 and 5 percentage points higher than the one of a typical firm with an asset composition ratio of 1.6 for all debt maturities.[28] The reason is that a higher asset composition ratio increases the default probability, particularly so in recessions where default losses are larger and harder to bear. Due to the resulting higher costs of debt, firms with growth options optimally select lower initial leverage.

As argued by Bhamra, Kuehn, and Strebulaev (2010a), however, it can be misleading to make quantitative statements simply based on optimal leverage at issue. Hence, we investigate the leverage ratios of our cross-section of BBB-rated firms simulated over 5 years after matching. For the main analysis, the debt maturity is assumed to be 10 years.

INSERT TABLE 5 HERE

Panel A in Table 5 shows that the average leverage is 40.89%, which is, naturally, close to the average of 41.83% of our BBB-rated firm sample used for the matching. (The average leverage is 40.57%, 40.93%, and 41.45% for 5 years, 20 years, and infinite debt maturity, respectively.)

[28] The difference depends on the initial regime and the debt maturity. For example, with infinite debt maturity, the difference in optimal initial leverage between a firm with only invested assets and a firm with an asset composition ratio of 1.6 is 4% if the firms are initiated in boom. (The optimal leverage ratios in this case are 45.4% and 41.4%, respectively.) For firms initiated in recession, the difference is 4.4% (= 44.2% minus 39.8%).

In Panel B, we compare leverage ratios in boom and recession. While optimal leverage is pro-cyclical at initiation, it is counter-cyclical over time for the cross-section of BBB-rated firms. In particular, the average leverage is 36.94% in boom, and 46.20% in recession. The reason is that the market value of equity is more sensitive to regime switches than the market value of debt, making leverage counter-cyclical. This mechanism dominates the optimally pro-cyclical leverage choice at initiation for our typical firms. The result mirrors the property we previously established for the aggregate economy, and confirms that it holds also when matching to real empirical samples.

Finally, Panel C investigates the relationship between growth options and market leverage. Regressing the average leverage of each firm on its average asset composition ratio in our empirical BBB-rated firm sample yields a coefficient of -0.165. We conduct the same regressions with the averages of asset composition ratios and leverage ratios from each of the 1000 simulations of the true cross-section. The average coefficient from this regression is -0.184, close to its empirical counterpart. Hence, the observed magnitude of the negative relationship between growth options and market leverage is preserved during the simulation.

Our qualitative finding for the cross-sectional relationship between growth options and leverage is widely accepted (Bradley, Jarrell, and Kim, 1984; Barclay, Smith, and Morellec, 2006; Johnson, 2003; Rajan and Zingales, 1995). Consistent with the literature, the coefficient is robustly negative. Moreover, its quantitative size, implied by the 25% and 75% quantiles, is comparable to the one in empirical studies. Fama and French (2002), for example, obtain a coefficient of -0.096 in their regression of market leverage on a similar ratio of asset composition after controlling for standard controls, and Johnson (2003) finds that increasing the asset composition ratio by one decreases leverage by around 7.8 percentage points in a pooled regression.

6.3. Robustness

In this section, we discuss the robustness of the results to variations in the critical input parameters. Additionally, we also show how our predictions are affected if we assume that the expansion is financed by issuing equity instead of selling assets.

To analyze the impact of preferences on our results, we show 10 year credit spreads and the simulated average leverage for $\gamma = 7.5$ in the second column of Table 6, a value which is also sometimes used in the literature (Bansal and Yaron, 2004; Chen, 2010). All other parameters are kept constant at their baseline levels from Table 1. The debt maturity is assumed to be 10 years.

INSERT TABLE 6 HERE

Lower risk aversion induces a smaller demand for precautionary savings, which increases the real risk-free rate. At the same time, it raises the risk-neutral earnings drift, because the risk prices for systematic Brownian shocks (η_i) decrease. Both mechanisms reduce the default probability, leading to the lower credit spreads and slightly lower leverage.

In column 3 of Table 6, we investigate the impact of the exercise costs on credit spreads and leverage. As we are mainly interested in firms with intact expansion options, we present the results for K equal to 350, i.e., a higher K than in the baseline case. (Lowering K induces many growth firms to exercise their expansion option almost immediately.) Generally, credit spreads and the average leverage are very similar to the ones of our baseline specification. For large asset composition ratios, such as with 2.2, credit spreads at initiation slightly increase because a higher K induces a larger distance to the optimal exercise boundary compared to the baseline specification. This increase in credit spreads from the larger distance arises because close to the exercise boundary, credit spreads reflect the fact that the firm will imminently be converted into a firm with only invested assets, and, hence, with lower credit risk.

Finally, we also analyze in Column 4 of Table 6 the case where the exercise price of the expansion option (K) is financed by issuing additional equity instead of selling assets. Appendix A.7.1 presents the resulting system of ODEs for corporate debt. New equity decreases the leverage after exercise and, hence, lowers credit risk. As firms with a high asset composition ratio are close to the endogenous exercise boundary where new equity-financing occurs, credit spreads are strongly reduced for typical growth firms compared to the benchmark model. In the simulation of the true cross-section, however, the effect is small because most firms have a large distance to the exercise boundary.[29] Additionally, Panel B shows that the average leverage is only marginally affected.

The result for typical growth firms in column 4 shows that close to firms' exercise boundaries, credit spreads are driven by the expected new financing upon investment, and do not primarily reflect the nature of assets. This insight validates our focus on asset-financing rather than on equity-financing of growth option exercises to analyze the isolated impact of the asset composition on credit risk and corporate policy choices.

We conclude that while alternative specifications and settings can have an impact on the quantitative results, our qualitative aggregate and cross-sectional predictions are robust.

7. Conclusion

It is now well-accepted that macroeconomic risk is central for understanding credit risk and capital structure choices. Specifically, defaults are more likely in recession, when they are particularly costly and harder to bear. This counter-cyclicality increases the costs of debt for all firms. But to explain the cross-sectional variation in apparently excessive costs of debt, we need variation inside

[29] In fact, those firms which contribute the most to the average credit spread, i.e., distressed firms, are particularly far away from the exercise boundary.

the firm. This paper formalizes the role of one particularly important aspect of this heterogeneity, the asset composition of firms. It is not surprising that in principle the asset composition can be important for optimal capital structure. After all, economists have devoted much effort to understanding the difference between value and growth firms in terms of their financial structure, starting with Myers (1977) and Jensen (1986). Little was known, however, about the quantitative importance of this factor and its relation with macroeconomic risk.

The present structural equilibrium model allows us to jointly analyze a firm's expansion policy and financial leverage in the presence of macroeconomic risk. We demonstrate that, in fact, incorporating the combination of these factors goes a long way towards explaining average credit spread levels, and the cross-sectional variation in both costs of debt and leverage without the need to appeal to factors such as agency costs. Our model implies that companies with a high portion of expansion options tend to be riskier in general, and, at the same time, particularly sensitive to macroeconomic risk. They are not only more volatile (because growth options represent levered claims), but also have a higher propensity to default in bad times than firms with a low portion of expansion options. Thus, the default probability and its counter-cyclicality are higher the greater the ratio of expansion options to total assets. Together with higher marginal utility of the representative agent in recession, this relation (exacerbated by costly liquidation in recession) implies higher costs of debt and more important endogenous shadow costs of leverage for firms with growth options than for those with only invested assets. Thus, our findings explain why the credit spread puzzle is empirically more pronounced for growth firms, and why growth firms hold less debt even after controlling for standard determinants of credit risk. Moreover, because the economy is made up of a cross-sectional mix of firms, the model accounts, in quantitatively fairly accurate ways, for the average credit spread puzzle.

We have studied one type of (arguably important) real options of firms, namely, growth options. However, firms have a wide and varying range of options, including abandonment and shut-down options. A model incorporating these options could, therefore, yield further cross-sectional predictions.

While recent research has made important progress in enhancing our understanding of average credit risk, the cross-section of credit risk has not received sufficient attention. Analyzing it empirically is, fortunately, quite feasible. Liquid credit default swap quotes are now widely available on a firm-by-firm basis, allowing researchers to investigate specific relationships between firm-specific characteristics such as growth options and credit spreads. Our paper also provides a theoretical basis that can guide empirical research in this direction.

References

Acharya, Viral, Sreedhar T. Bharath, and Anand Srinivasan, 2007, Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries, *Journal of Financial Economics* 85, 787–821.

Almeida, Heitor, and Thomas Philippon, 2007, The risk-adjusted cost of financial distress, *Journal of Finance* 62, 2557–2586.

Ang, Andrew, and Geert Bekaert, 2004, How regimes affect asset allocation, *Financial Analysts Journal* 60, 86–99.

Baker, Malcolm, and Jeffrey Wurgler, 2002, Market timing and capital structure, *Journal of Finance* 57, 1–32.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1507.

Barclay, Michael J., Clifford W. Smith, and Erwan Morellec, 2006, On the debt capacity of growth options, *Journal of Business* 79, 37–59.

Barro, Robert J., 1990, The stock market and investment, *Review of Financial Studies* 3, 115–131.

Berk, Jonathan B., C. Green, Richard, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Bhagat, Sanjai, Andrei Shleifer, and Robert W. Vishny, 1990, Hostile takeovers in the 1980s: The return to corporate specialization, *Brookings Papers on Economic Activity: Microeconomics* pp. 1–72.

Bhamra, Harjoat S., Lars-Alexander Kuehn, and Ilya A. Strebulaev, 2010a, The aggregate dynamics of capital structure and macroeconomic risk, *Review of Financial Studies* 23, 4187–4241.

——— , 2010b, The levered equity risk premium and credit spreads: A unified framework, *Review of Financial Studies* 23, 645–703.

Bonomo, Marco, and Ren Garcia, 1996, Consumption and equilibrium asset pricing: An empirical assessment, *Journal of Empirical Finance* 3, 239–265.

Bradley, Michael, Gregg A. Jarrell, and Han E. Kim, 1984, On the existence of an optimal capital structure: Theory and evidence, *Journal of Finance* 39, 857–878.

Cantor, Richard, and Praveen Varma, 2005, Determinants of recovery rates on defaulted bonds and loans for North American corporate issuers: 1983-2003, *Journal of Fixed Income* 14, 29–44.

Carlson, Murray D., Adlai J. Fisher, and Ron Giammarino, 2006, Corporate investment and asset price dynamics: Implications for SEO event studies and long-run performance, *Journal of Finance* 61, 1009–1034.

Chava, Sudheer, and Robert A. Jarrow, 2004, Bankruptcy prediction with industry effects, *Review of Finance* 8, 537–569.

Chen, Hui, 2010, Macroeconomic conditions and the puzzles of credit spreads and capital structure, *Journal of Finance* 65, 2171–2212.

——— , and Gustavo Manso, 2010, Macroeconomic risk and debt overhang, *Mimeo.*

Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein, 2009, On the relation between the credit spread puzzle and the equity premium puzzle, *Review of Financial Studies* 22, 3367–3409.

Childs, Paul, D., David C. Mauer, and Steven H. Ott, 2005, Interactions of corporate financing and investment decisions: The effects of agency conflicts, *Journal of Financial Economics* 76, 667–690.

Collin-Dufresne, Pierre, and Robert S. Goldstein, 2001, Do credit spreads reflect stationary leverage ratios?, *Journal of Finance* 56, 1929–1957.

Cooper, Russell, John Haltiwanger, and Laura Power, 1999, Machine replacement and the business cycle: Lumps and bumps, *American Economic Review* 89, 921–946.

Covas, Francisco, and Wouter J. Den Haan, 2006, The role of debt and equity finance over the business cycle, *Mimeo.*

Daines, Rober M., Ian D. Gow, and David F. Larcker, 2010, Rating the ratings: How good are commercial governance ratings?, *Journal of Financial Economics* 98, 439–461.

Das, Sanjiv R., Darrell Duffie, Nikunj Kapadia, and Leandro Saita, 2007, Common failings: How corporate defaults are correlated, *Journal of Finance* 62, 93–117.

Datta, Sudip, Mai Iskandar-Datta, and Ajay Patel, 2000, Some evidence on the uniqueness of initial public debt offerings, *Journal of Finance* 55, 715–743.

Davydenko, Sergei A., and Ilya A. Strebulaev, 2007, Strategic actions and credit spreads: An empirical investigation, *Journal of Finance* 62, 2633–2671.

De Jong, Frank, and Driessen Joost, 2006, Liquidity risk premia in corporate bond markets, *Mimeo.*

Demchuk, Andriy, and Rajna Gibson, 2006, Stock market performance and the term structure of credit spreads, *Journal of Financial and Quantitative Analysis* 41, 863–887.

Doms, Mark E., and Timothy Dunne, 1998, Capital adjustment patterns in manufacturing plants, *Review of Economic Dynamics* 1, 402–429.

Duan, Jin-Chuan, 1994, Maximum likelihood estimation using price data of the derivatives contract, *Mathematical Finance and Stochastics* 4, 155–167.

Duffee, Gregory R., 1998, The relation between treasury yields and corporate bond yield spreads, *Journal of Finance* 53, 2225–2241.

Duffie, Darrell, and Larry G. Epstein, 1992a, Asset pricing with stochastic differential utility, *Review of Financial Studies* 5, 411–436.

——— , 1992b, Stochastic differential utility, *Econometrica* 60, 353–394.

Duffie, Darrell, Leandro Saita, and Ke Wang, 2007, Multi-period corporate default prediction with stochastic covariates, *Journal of Financial Economics* 83, 635–665.

Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, and Christopher Mann, 2001, Explaining the rate spread on corporate bonds, *Journal of Finance* 56, 247–277.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–69.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

——— , 2002, Testing trade-off and pecking order predictions about dividends and debt, *Review of Financial Studies* 15, 1–33.

Frank, Murray Z., and Vidhan K. Goyal, 2009, Capital structure decisions: Which factors are reliably important?, *Financial Management* 38, 1–37.

Frye, Jon, 2000, Collateral damage, *Risk* April, 91–94.

Gopalan, Radhakrishnan, Fenghua Song, and Vijay Yerramilli, 2010, Debt maturity structure and credit quality, *Mimeo.*

Graham, John R., 2000, How big are the tax benefits of debt?, *Journal of Finance* 53, 1901–1941.

Guedes, Jose, and Timothy Opler, 1996, The determinants of the maturity of corporate debt issues, *Journal of Finance* 51, 1809–1833.

Guo, Xin, 2001, An explicit solution to an optimal stopping problem with regime switching, *Jounal of Applied Probability* 38, 464–481.

——— , and Qing Zhang, 2004, Closed-form solutions for perpetual American put options with regime switching, *Journal of Applied Mathematics* 64, 2034–2049.

Hackbarth, Dirk, and David C. Mauer, 2010, Optimal priority structure, capital structure, and investment, *Mimeo.*

Hackbarth, Dirk, Jianjun Miao, and Erwan Morellec, 2006, Capital structure, credit risk, and macroeconomic conditions, *Journal of Financial Economics* 82, 519–550.

Han, Song, and Hao Zhou, 2011, Effects of liquidity on the nondefault component of corporate yield spreads: Evidence from intraday transactions data, *Mimeo.*

He, Zhiguo, and Wei Xiong, 2011, Rollover risk and credit risk, *Mimeo.*

Huang, Jing-Zhi, and Ming Huang, 2003, How much of the corporate-treasury yield spread is due to credit risk? A new calibration approach, *Mimeo.*

Jensen, Michael C., 1986, Agency costs of free cash flow, corporate finance and takeovers, *American Economic Review* 76, 323–329.

Johnson, Shane A., 2003, Debt maturity and the effects of growth opportunities and liquidity risk on leverage, *Review of Financial Studies* 16, 209–236.

Kaplan, Steven N., and Michael S. Weisbach, 1992, The success of acquisitions: Evidence from divestitures, *Journal of Finance* 47, 107–138.

Korajczyk, Robert A., and Amnon Levy, 2003, Capital structure choice: Macroeconomic conditions and financial constraints., *Journal of Financial Economics* 68, 75–109.

Korteweg, Arthur G., 2011, The net benefits to leverage, *Journal of Finance, forthcoming.*

Leland, Hayne E., 1994, Corporate debt value, bond covenants, and optimal capital structure, *Journal of Finance* 49, 1213–1252.

——— , 1998, Agency costs, risk management, and capital structure, *Journal of Finance* 53, 1213–1243.

Longstaff, Francis A., Sanjay Mithal, and Eric Neis, 2005, Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market, *Journal of Finance* 60, 2213–2253.

Lucas, Robert E., 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1445.

Lyandres, Evgeny, and Alex Zhdanov, 2010, Accelerated investment effect of risky debt, *Journal of Banking and Finance* 34, 2587–2599.

Mello, Antonio, and John Parsons, 1992, Measuring the agency cost of debt, *Journal of Finance* 47, 1887–1904.

Miao, Jianjun, 2005, Optimal capital structure and industry dynamics, *Journal of Finance* 60, 2621–2659.

Molina, Carlos A., 2005, Are firms underleveraged? An examination of the effect of leverage on default probabilities, *Journal of Finance* 60, 1427–1459.

Moody's, 2010, Corporate default and recovery rates 1920-2009, *Mimeo.*

Morellec, Erwan, 2004, Can managerial discretion explain observed leverage ratios?, *Review of Financial Studies* 17, 257–294.

——— , Boris Nikolov, and Norman Schürhoff, 2009, Dynamic capital structure under managerial entrenchment: Evidence from a structural estimation, *Mimeo.*

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.

Oivind, Anti N., and Fabio Schiantarelli, 2003, Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities, *Review of Economics and Statistics* 85, 1021–1037.

Parrino, Robert, and Michael S. Weisbach, 1999, Measuring investment distortions arising from stockholder-bondholder conflicts, *Journal of Financial Economics* 53, 3–42.

Polyanin, Andrei D., and Valentin R. Zaitsev, 2003, *Exact solutions for ordinary differential equations* (Chapman & Hall /CRC) 2 edn.

Rajan, Raghuram G., and Luigi Zingales, 1995, What do we know about capital structure? Some evidence from international data, *Journal of Finance* 50, 1421–1467.

Schaefer, Stephen M., and Ilya A. Strebulaev, 2008, Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds, *Journal of Financial Economics* 90, 1–19.

Shleifer, Andrei, and Robert W. Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance* 47, 1343–1266.

Smith, Clifford Jr., and Ross L. Watts, 1992, The investment opportunity set and corporate financing, dividend, and compensation policies, *Journal of Financial Economics* 32, 262–292.

Strebulaev, Ilya A., 2007, Do tests of capital structure theory mean what they say?, *Journal of Finance* 62, 1747–1787.

Sundaresan, Suresh, and Heng Wang, 2007, Dynamic investment, capital structure, and debt overhang, *Mimeo.*

Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831–868.

Weil, Philippe, 1990, Nonexpected utility in macroeconomics, *Quarterly Journal of Economics* 105, 29–42.

Yin, George, Qingshue Song, and Zhimin Zhang, 2004, Numerical solutions for jump-diffusions with regime switching, *Stochastics* 77, 61–79.

8. Figures

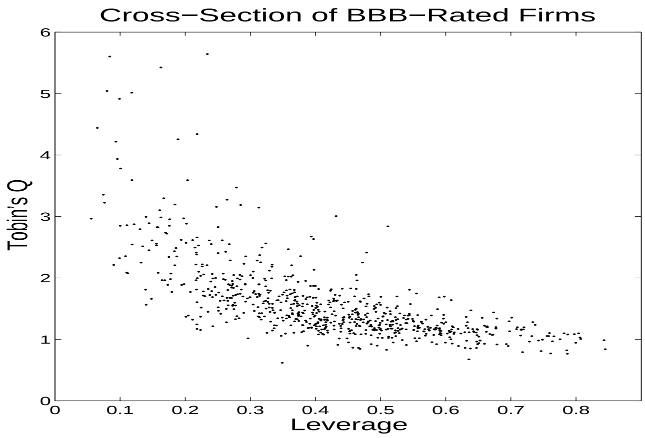

Figure 1. This scatterplot shows the average leverage and Tobin's Q for each observed BBB-rated firm over the period from 1995 to 2008.

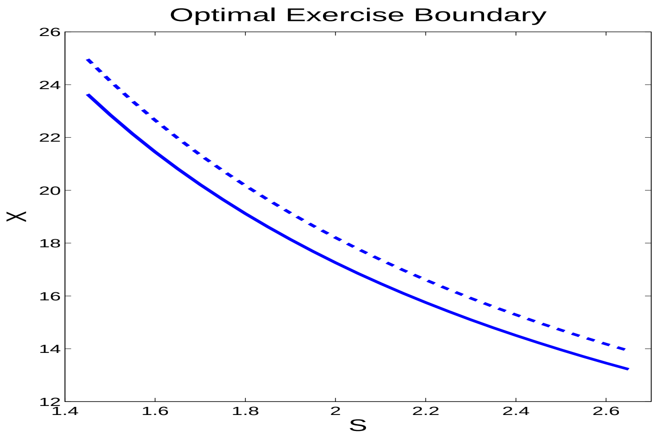

Figure 2. The solid line shows the optimal exercise boundary in boom for a range of scale parameters s. The dashed line represents the corresponding exercise boundary in recession. The graph is drawn for optimal leverage with infinite debt maturity. The baseline parameter specification from Table 1 is used.

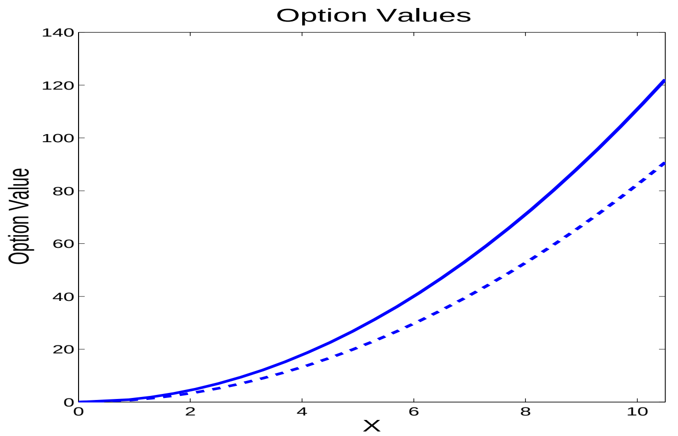

Figure 3. The solid line represents the value of the expansion option in boom for a range of starting earnings between 0 and 10. The dashed line shows the corresponding values of the same option in recession. The graph is drawn for optimal leverage with infinite debt maturity. The baseline parameter specification from Table 1 is used.

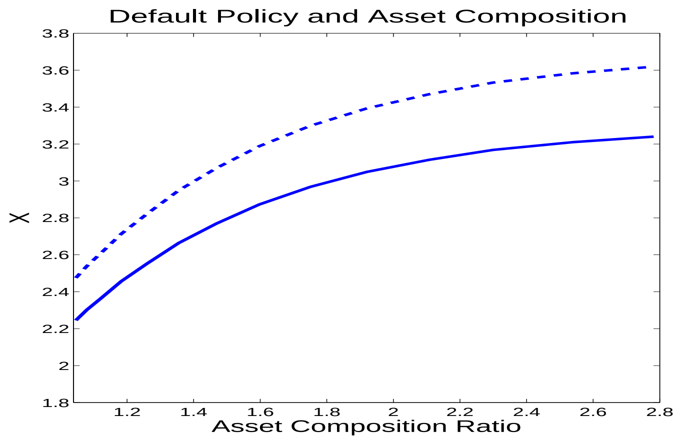

Figure 4. The solid line represents the default threshold in boom for a range of asset composition ratios. The dashed line shows the default threshold in recession. The graph is drawn for constant leverage (41.83%) at each point. Debt maturity is assumed to be infinite. The baseline parameter specification from Table 1 is used, with s being varied to generate the desired asset composition ratio.

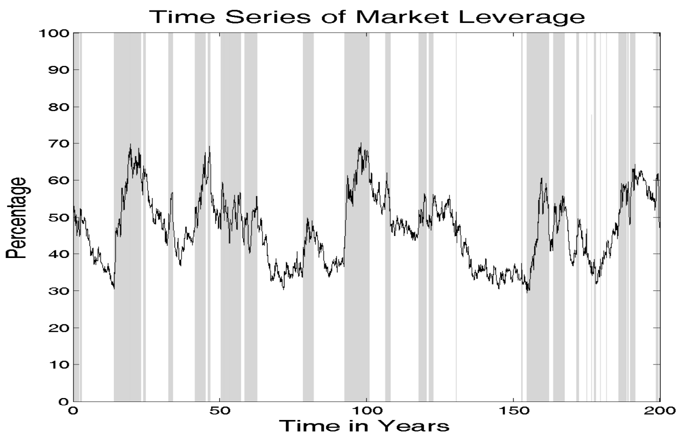

Figure 5. The solid line represents the aggregate market leverage of the simulated economy. The shaded areas represent times of recession. Standard parameters from Table 1 are used. Debt maturity is assumed to be infinite.

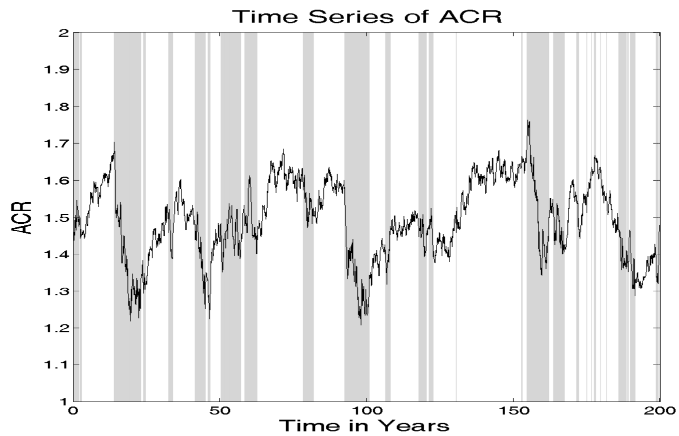

Figure 6. The solid line represents the aggregate asset composition ratio of the simulated economy. The shaded areas represent times of recession. Standard parameters from Table 1 are used. Debt maturity is assumed to be infinite.

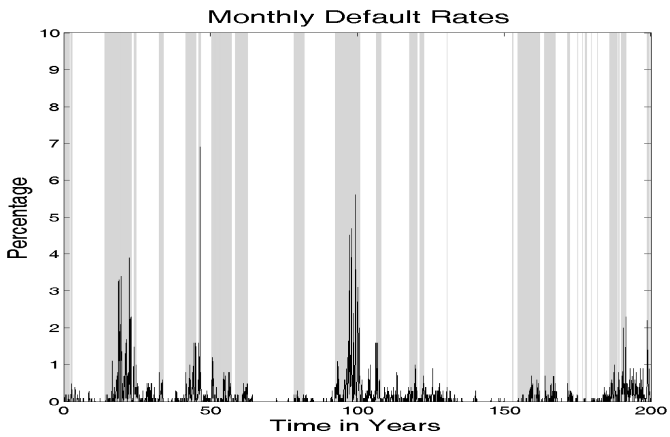

Figure 7. The solid line represents the percentage of firms which default during a given month in the simulated economy. The shaded areas represent times of recession. Standard parameters from Table 1 are used. Debt maturity is assumed to be infinite.

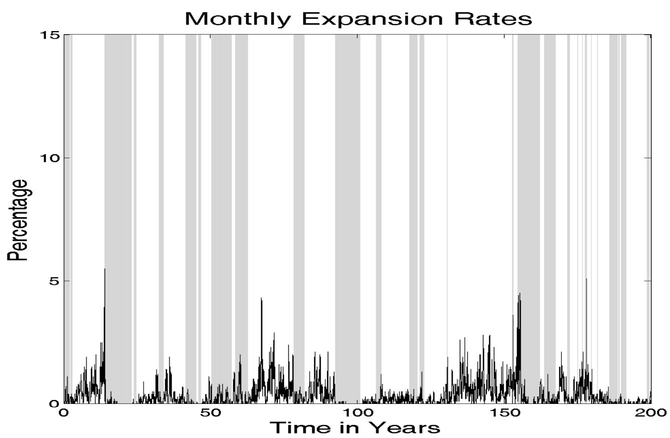

Figure 8. The solid line represents the percentage of firms which exercise their expansion options during a given month in the simulated economy. The shaded areas represent times of recession. Standard parameters from Table 1 are used. Debt maturity is assumed to be infinite.

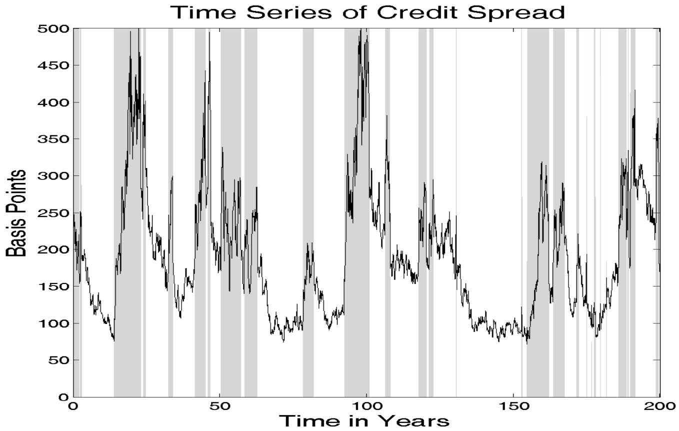

Figure 9. The solid line represents the average credit spread of the simulated economy. The shaded areas represent times of recession. Standard parameters from Table 1 are used. Debt maturity is assumed to be infinite.

9. Tables

Table 1
Baseline Parameter Choice

This table describes our baseline scenario. Panel A contains the calibrated parameters of a typical BBB-rated S&P 500 firm. Panels B and C show our parameter choice for the expansion option and our workhorse macro economy, respectively. The asset composition ratio (ACR) is the value of the firm, divided by the value of the invested assets.

Parameter	Boom	Recession
Panel A. Firm Characteristics		
Initial Value of Idiosyncratic Earnings (X)	10	10
Tax Advantage of Debt (τ)	0.15	0.15
Real Earnings Growth Rate (μ_i)	0.044	-0.0743
Systematic Earnings Volatility ($\sigma_i^{X,C}$)	0.0869	0.1369
Recovery Rate (α_i)	0.7	0.5
Panel B. Expansion Option Parameters of a Typical Firm (ACR=1.6)		
Exercise Price (K)	310	310
Scale Parameter if Initiated in Boom (s)	1.89	
Scale parameter if Initiated in Recession (s)	2.05	
Panel C. Economy		
Regime Switching Intensity (λ)	0.2718	0.4928
Consumption Growth Rate (θ_i)	0.042	0.0141
Consumption Growth Volatility (σ_i^C)	0.0094	0.0114
Expected Inflation Rate (π)	0.0342	0.0342
Systematic Price Index Volatility ($\sigma^{P,C}$)	-0.0035	-0.0035
Idiosyncratic Price Index Volatility ($\sigma^{P,id}$)	0.0132	0.0132
Rate of time preference (ρ)	0.015	0.015
Relative Risk Aversion (γ)	10	10
Elasticity of Intertemporal Substitution (Ψ)	1.5	1.5

Table 2

Target Credit Spreads and Default Probabilities

This table lists our target credit spreads and default probabilities. Panel A reports target average credit spreads for various debt maturities. They are calculated as the BBB-rated bond minus treasury yields of Davydenko and Strebulaev (2007) and Duffee (1998), net of a 35.5% non-default component. Credit spreads are quoted in basis points. Panel B reports average cumulative issuer-weighted default rates in percent for BBB-debt over 5, 10, and 20 years for US firms (Moody's, 2010).

Panel A: Target Credit Spreads (in basis points)			
Debt Maturity	Short	Medium	Long
Davydenko and Strebulaev (2007)	74	74	92
Duffee (1998)	96	95	128
Panel B: Historical BBB Default Probabilities (in percent)			
Years	5	10	20
1920-2009	3.136	7.213	13.684
1970-2009	1.926	4.851	12.327

Table 3
Implications for Credit Spreads

This table demonstrates the implications of our model for credit spreads of BBB-rated firms. The asset composition ratio (ACR) is defined as firm value, divided by the value of the invested assets. Parameters are taken from Table 1, and the leverage is set equal to 41.83%. In the one regime model, parameters are chosen to match their unconditional mean. The standard two regime model is adapted from Bhamra, Kuehn, and Strebulaev (2010b). Credit spreads for various debt maturities are calculated as the coupon divided by the debt value, minus the yield on an otherwise identical riskfree bond. They are quoted in basis points. Credit spreads of typical firms in Panels B and C are obtained by weighting the credit spreads in boom and recession by the average expected times spent in each regime, respectively. Panel D contains the average credit spreads of our simulated true cross-section of BBB-rated firms.

Debt Maturity (Years)	5	10	20
Panel A: One Regime Model			
Average Firm	18	29	41
Panel B: Standard Two Regime Model With Only Invested Assets			
Average Firm	35	56	78
Panel C: Two Regime Model With Expansion Option			
Invested Assets (ACR=1)	24	39	55
Average Firm (ACR=1.6)	44	66	81
Growth Firm (ACR=2.2)	47	70	85
Panel D: Two Regime Model With True Cross-Section			
Average Credit Spread	60	81	103

Table 4

Implications for Default Rates

This table shows the simulated cumulative default rates in percent of our true cross-section of BBB-rated firms. Panels A to D vary the underlying debt maturity used to calculate the default thresholds in our model.

Years	5	10	20
Panel A: Infinite Debt Maturity			
Average Default Rates	2.51	6.94	13.48
Median Default Rates	0.98	3.35	9.21
25% Quantile of Default Rates	0.35	1.26	3.63
75% Quantile of Default Rates	2.65	8.79	18.69
Panel B: 20 Years Debt Maturity			
Average Default Rates	4.51	10.58	18.61
Median Default Rates	1.95	5.44	13.11
25% Quantile of Default Rates	0.70	2.09	5.44
75% Quantile of Default Rates	5.30	13.95	26.08
Panel C: 10 Years Debt Maturity			
Average Default Rates	5.85	12.38	20.54
Median Default Rates	2.65	6.83	14.37
25% Quantile of Default Rates	0.98	2.93	6.83
75% Quantile of Default Rates	6.83	16.88	30.40
Panel D: 5 Years Debt Maturity			
Average Default Rates	8.64	16.91	25.96
Median Default Rates	4.74	11.92	20.36
25% Quantile of Default Rates	1.81	5.02	10.18
75% Quantile of Default Rates	11.02	24.13	38.49

Table 5

Implications for Leverage

This table demonstrates the implications of our model for the leverage features of the true cross-section of BBB-rated firms. Leverage ratios (given in percent) are calculated as the market value of debt divided by the market value of the firm. The asset composition ratio (ACR) is defined as firm value, divided by the value of the invested assets. Parameters are taken from Table 1. The debt maturity is assumed to be 10 years.

Panel A: Unconditional Leverage		
Average Leverage	40.89	
Panel B: Conditional Leverage		
Regime	Boom	Recession
Average Leverage	36.94	46.20
Median Leverage	34.36	44.19
25% Quantile	22.49	29.88
75% Quantile	48.51	60.39
Panel C: Regression of Leverage on ACR		
Average Coefficient	-0.184	
Median Coefficient	-0.184	
25% Quantile	-0.268	
75% Quantile	-0.096	

Table 6

Credit Spreads and Leverage for Alternative Specifications

This table shows 10 year credit spreads and simulated average leverage ratios of BBB-rated firms for alternative specifications of our basic model. The asset composition ratio (ACR) is defined as firm value, divided by the value of the invested assets. Credit spreads are calculated as the coupon divided by the debt value, minus the yield on an otherwise identical riskfree bond. They are quoted in basis points. The altered parameter is indicated in the first line, all other parameters are taken from Table 1. Credit spreads in the first 3 lines of Panel A for typical firms at issue are obtained by weighting the credit spreads in boom and recession by the expected times spent in each regime, respectively. The leverage is set equal to 41.83% to generate the credit spreads of typical firms. The last row in Panel A contains average credit spreads of our simulated true cross-section of BBB-rated firms. Panel B shows simulated average leverage ratios for BBB-rated firms. The debt maturity is assumed to be 10 years.

Specification	$\gamma = 7.5$	$K = 350$	Equity Financing
Panel A: 10 Year Credit Spreads			
Invested Assets (ACR=1)	33	39	39
Average Firm (ACR=1.6)	53	67	65
Growth Firm (ACR=2.2)	56	72	58
True Cross-Section	70	83	80
Panel B: Unconditional Leverage			
Average Leverage	41.10	41.22	41.14

A. Appendix

The full Appendix can be made available on a website upon publication.

A.1. The stochastic discount factor

Case 1: The general case with 2 regimes. Solving the associated Bellman equation (see Chen, 2010), it can be shown that the stochastic discount factor m_t follows the dynamics

$$\frac{\mathrm{d}m_t}{m_t} = -r_i \mathrm{d}t - \eta_i \mathrm{d}W_t^C + \left(e^{\kappa_i} - 1\right) \mathrm{d}M_t^i, \tag{A-1}$$

with M_t being the compensated process associated with the Markov chain, and

$$r_i = -\rho \frac{(1-\gamma)}{1-\delta} \left(\frac{\delta - \gamma}{1-\gamma} h_i^{\delta-1} - 1 \right) + \gamma \theta_i - \frac{1}{2} \gamma (1+\gamma) \left(\sigma_i^C\right)^2 - \lambda_i \left(e^{\kappa_i} - 1\right) \tag{A-2}$$

$$\eta_i = \gamma \sigma_i^C \tag{A-3}$$

$$\kappa_i = (\delta - \gamma) \log \left(\frac{h_j}{h_i} \right), \tag{A-4}$$

and h_B, h_R solve

$$0 = \rho \frac{1-\gamma}{1-\delta} h_i^{\delta-\gamma} + \left((1-\gamma) \theta_i - \frac{1}{2} \gamma (1-\gamma) \left(\sigma_i^C\right)^2 - \rho \frac{1-\gamma}{1-\delta} \right) h_i^{1-\gamma} + \lambda_i \left(h_j^{1-\gamma} - h_i^{1-\gamma} \right). \tag{A-5}$$

Case 2: Only 1 regime. In order to disentangle the effect of business cycle risk, we also consider the case of the presence of only one economic regime. We omit regime indeces. The dynamics of the stochastic discount factor then read

$$\frac{\mathrm{d}m_t}{m_t} = -r \mathrm{d}t - \eta \mathrm{d}W_t^C. \tag{A-6}$$

The real interest rate r and the risk price η are given by

$$r = -\frac{\rho (1-\gamma)}{1-\delta} \left(\frac{\delta - \gamma}{1-\gamma} h^{\delta-1} - 1 \right) + \gamma \theta - \frac{1}{2} \gamma (1+\gamma) \left(\sigma^C\right)^2, \tag{A-7}$$

$$\eta = \gamma \sigma^C, \tag{A-8}$$

with

$$h = \left. - \frac{\rho}{(1-\delta) \theta - \frac{1}{2} \gamma (1-\delta) \left(\sigma^C\right)^2 - \rho} \right)^{\frac{1}{1-\delta}}. \tag{A-9}$$

As before, the nominal interest rate is calculated as

$$r^n = r + \pi - \sigma_P^2 - \sigma^{P,C} \eta, \tag{A-10}$$

and the expected growth rate is given by

$$\tilde{\mu} = \mu - \sigma^{X,C} \left(\eta + \sigma^{P,C} \right) - \left(\sigma^{P,id} \right)^2. \tag{A-11}$$

The price-earnings ratio simplifies to

$$y = \frac{1}{r^n - \tilde{\mu}}, \tag{A-12}$$

and the total earnings volatility is

$$\tilde{\sigma} = \sqrt{\left(\sigma^{X,C}\right)^2 + \left(\sigma^{P,id}\right)^2 + \left(\sigma^{X,id}\right)^2}. \tag{A-13}$$

A.2. Firms with only invested assets

A.2.1. The valuation of corporate debt

Case V1: $\hat{D}_B < \hat{D}_R$.[30] We use the notation '$\hat{\ }$' to indicate that a parameter or function refers to a firm with only invested assets (e.g. the default boundaries $\hat{D}_i$). An investor holding corporate debt requires an instantaneous return equal to the risk-free rate r_i^n. Once the firm defaults, debtholders receive a fraction α_i of the asset value Xy_i. The required rate of return on debt must be equal to the realized rate of return plus the proceeds of debt. Therefore, an application of Ito's lemma with regime switches shows that debt satisfies the following system of ODEs:
For $0 \leq X \leq \hat{D}_B$:

$$\begin{cases} \hat{d}_B(X) &= \alpha_B X y_B \\ \hat{d}_R(X) &= \alpha_R X y_R. \end{cases} \tag{A-14}$$

For $\hat{D}_B < X \leq \hat{D}_R$:

$$\begin{cases} r_B^n \hat{d}_B(X) &= c + \tilde{\mu}_B X \hat{d}_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 \hat{d}_B''(X) + \tilde{\lambda}_B \left(\alpha_R X y_R - \hat{d}_B(X)\right) \\ \hat{d}_R(X) &= \alpha_R X y_R. \end{cases} \tag{A-15}$$

For $X > \hat{D}_R$:

$$\begin{cases} r_B^n \hat{d}_B(X) &= c + \tilde{\mu}_B X \hat{d}_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 \hat{d}_B''(X) + \tilde{\lambda}_B \left(\hat{d}_R(X) - \hat{d}_B(X)\right) \\ r_R^n \hat{d}_R(X) &= c + \tilde{\mu}_R X \hat{d}_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 \hat{d}_R''(X) + \tilde{\lambda}_R \left(\hat{d}_B(X) - \hat{d}_R(X)\right). \end{cases} \tag{A-16}$$

The functional form of the solution is

$$\hat{d}_i(X) = \begin{cases} \alpha_i X y_i & X \leq \hat{D}_i & i = B, R \\ \hat{C}_1 X^{\beta_1^B} + \hat{C}_2 X^{\beta_2^B} + C_3 X + C_4 & \hat{D}_B < X \leq \hat{D}_R, & i = B \\ \hat{A}_{i1} X^{\gamma_1} + \hat{A}_{i2} X^{\gamma_2} + A_{i5} & X > \hat{D}_R, & i = B, R, \end{cases} \tag{A-17}$$

where $\hat{A}_{B1}, \hat{A}_{B2}, \hat{A}_{R1}, \hat{A}_{R2}, A_5, \hat{C}_1, \hat{C}_2, C_3, C_4, \gamma_1, \gamma_2, \beta_1^B$, and β_2^B are real-valued parameters to be determined. We first consider the region $X > \hat{D}_R$, and use the standard approach of plugging in the functional

[30]The solution of the case $\hat{D}_B > \hat{D}_R$ can be found by the according change in notation.

form $\hat{d}_i(X) = \hat{A}_{i1}X^{\gamma_1} + \hat{A}_{i2}X^{\gamma_2} + A_{i5}$ into both equations of (A-16). Comparing coefficients and solving the resulting 2-dimensional system of equations for A_{i5}, we find that

$$A_{i5} = \frac{c\left(r_j^n + \tilde{\lambda}_i + \tilde{\lambda}_j\right)}{r_i^n r_j^n + r_j^n\tilde{\lambda}_i + r_i^n\tilde{\lambda}_j}. \tag{A-18}$$

Next, $\hat{A}_{Rk}$ is always a multiple of $\hat{A}_{Bk}$, $k = 1, 2$, with the factor $l_k := \frac{1}{\tilde{\lambda}_B}(r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B\gamma_k - \frac{1}{2}\tilde{\sigma}_B^2\gamma_k(\gamma_k - 1))$, i.e., $A_{Bk} = l_k A_{Rk}$.

Using these results and comparing coefficients again, we find that γ_1 and γ_2 correspond to the negative roots of the quartic equation

$$\left(\tilde{\mu}_R\gamma + \frac{1}{2}\tilde{\sigma}_R^2\gamma(\gamma-1) - \tilde{\lambda}_R - r_R^n\right)\left(\tilde{\mu}_B\gamma + \frac{1}{2}\tilde{\sigma}_B^2\gamma(\gamma-1) - \tilde{\lambda}_B - r_B^n\right) = \tilde{\lambda}_R\tilde{\lambda}_B, \tag{A-19}$$

with the reason for taking the negative roots being the no-bubbles condition for debt stated below. By arguments of Guo (2001), this quartic equation always has four distinct real roots, two of them being negative, and two of them positive.

Next, we consider the region $\hat{D}_B \le X \le \hat{D}_R$, i.e., the realized state of the Markov chain is boom (if not, the solution is already known by the second equation of system (A-15)). Again, plugging in the functional form $d_B(X) = \hat{C}_1 X^{\beta_1^B} + \hat{C}_2 X^{\beta_2^B} + C_3 X + C_4$ into the first equation of (A-15), we find by comparison of coefficients that

$$\begin{aligned} \beta_{1,2}^B &= \frac{1}{2} - \frac{\tilde{\mu}_B}{\tilde{\sigma}_B^2} \pm \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}_B}{\tilde{\sigma}_B^2}\right)^2 + \frac{2\left(r_B^n + \tilde{\lambda}_B\right)}{\tilde{\sigma}_B^2}} \\ C_3 &= \frac{\tilde{\lambda}_B \alpha_R y_R}{r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B} \\ C_4 &= \frac{c}{r_B^n + \tilde{\lambda}_B}, \end{aligned} \tag{A-20}$$

for $r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B \neq 0$. The unknown parameters are now $\hat{A}_{B1}, \hat{A}_{B2}, \hat{C}_1$ and $\hat{C}_2$. The boundary conditions read

$$\lim_{X\to\infty} \frac{\hat{d}_i(X)}{X} < \infty, \quad i = B, R \tag{A-21}$$

$$\lim_{X\searrow\hat{D}_R} \hat{d}_B(X) = \lim_{X\nearrow\hat{D}_R} \hat{d}_B(X) \tag{A-22}$$

$$\lim_{X\searrow\hat{D}_R} \hat{d}'_B(X) = \lim_{X\nearrow\hat{D}_R} \hat{d}'_B(X) \tag{A-23}$$

$$\lim_{X\searrow\hat{D}_B} \hat{d}_B(X) = \alpha_B D_B y_B \tag{A-24}$$

$$\lim_{X\searrow\hat{D}_B} \hat{d}_R(X) = \alpha_R D_R y_R. \tag{A-25}$$

Condition (A-21) is the no-bubbles condition used above in determining the appropriate roots of equation (A-19). The default thresholds $\hat{D}_R$ and $\hat{D}_B$ are chosen by the equityholders, and are taken as given by the debtholders. The boundary conditions are, hence, the value-matching conditions (A-22), (A-24), and (A-25), and the smooth-pasting condition at the higher default threshold $\hat{D}_B$ for the debt function in recession $\hat{d}_R(\cdot)$, equation (A-23). As the default thresholds are not related to an optimality concept from the point of view of the debtholders, there are no smooth-pasting conditions at default to consider.

We plug in the functional form (A-17) into conditions (A-22)-(A-25), and obtain a four-dimensional linear system in the four unknowns $\hat{A}_{B1}, \hat{A}_{B2}, \hat{C}_1$ and $\hat{C}_2$:

$$\begin{array}{rcl} \hat{A}_{B1}\hat{D}_R^{\gamma_1} + \hat{A}_{B2}\hat{D}_R^{\gamma_2} + A_5 & = & \hat{C}_1\hat{D}_R^{\beta_1^B} + \hat{C}_2\hat{D}_R^{\beta_2^B} + C_3\hat{D}_R + C_4 \\ \hat{A}_{B1}\gamma_1\hat{D}_R^{\gamma_1} + \hat{A}_{B2}\gamma_2\hat{D}_R^{\gamma_2} & = & \hat{C}_1\beta_1^B\hat{D}_R^{\beta_1^B} + \hat{C}_2\beta_2^B\hat{D}_R^{\beta_2^B} + C_3\hat{D}_R \\ \alpha_B\hat{D}_B y_B & = & \hat{C}_1\hat{D}_R^{\beta_1^B} + \hat{C}_2\hat{D}_R^{\beta_2^B} + C_3\hat{D}_R + C_4 \\ l_1\hat{A}_{B1}\hat{D}_R^{\gamma_1} + l_2\hat{A}_{B2}\hat{D}_R^{\gamma_2} + A_5 & = & \alpha_R\hat{D}_R y_R. \end{array} \tag{A-26}$$

We define the matrices

$$\begin{array}{rcl} \hat{M} & := & \begin{bmatrix} \hat{D}_R^{\gamma_1} & \hat{D}_R^{\gamma_2} & -\hat{D}_R^{\beta_1^B} & -\hat{D}_R^{\beta_2^B} \\ \gamma_1\hat{D}_R^{\gamma_1} & \gamma_2\hat{D}_R^{\gamma_2} & -\beta_1^B\hat{D}_R^{\beta_1^B} & -\beta_2^B\hat{D}_R^{\beta_2^B} \\ 0 & 0 & \hat{D}_R^{\beta_1^B} & \hat{D}_R^{\beta_2^B} \\ l_1\hat{D}_R^{\gamma_1} & l_2\hat{D}_R^{\gamma_2} & 0 & 0 \end{bmatrix} \\ \hat{b} & := & \begin{bmatrix} C_3\hat{D}_R + C_4 - A_{B5} \\ C_3\hat{D}_R \\ \alpha_B\hat{D}_B y_B - C_3\hat{D}_R - C_4 \\ \alpha_R\hat{D}_R y_R - A_{R5} \end{bmatrix}, \end{array}$$

such that $\hat{M}\begin{bmatrix}\hat{A}_{B1} & \hat{A}_{B2} & \hat{C}_1 & \hat{C}_2\end{bmatrix}^T = \hat{b}$. Hence the solution of the unknowns left is given by

$$\begin{bmatrix}\hat{A}_{B1} & \hat{A}_{B2} & \hat{C}_1 & \hat{C}_2\end{bmatrix}^T = \hat{M}^{-1}\hat{b}. \tag{A-27}$$

Case 2: Only 1 regime. Omitting the regime-index, we define all parameters and functions as in Case V1, and let $\hat{D}_1$ be the default threshold. Note that for $\mathbb{P} = \mathbb{Q}$ a.e. and $y = 1$, this case corresponds to the model of Leland (1994). Equations (A-10)-(A-13) provide all the parameters needed in the setup and solution of the model in the 1-regime case. Using that the required return must be equal to the expected realized return plus the proceeds from debt, we find the following system to solve:

$$\begin{array}{rcll} r^n\hat{d}(X) & = & c + \tilde{\mu}X\hat{d}'(X) + \frac{\tilde{\sigma}^2}{2}X^2\hat{d}''(X) & X > \hat{D} \\ \hat{d}(X) & = & \alpha X y & X \leq \hat{D}. \end{array} \tag{A-28}$$

The boundary conditions are the no bubbles condition, as well as value-matching at default:

$$\begin{array}{rcl} \lim\limits_{X\to\infty} \dfrac{\hat{d}(X)}{X} & < & \infty \\ \lim\limits_{X\searrow\hat{D}} \hat{d}\,(X) & = & \alpha y\hat{D}. \end{array} \tag{A-29}$$

The functional form of the solution is

$$\hat{d}(X) = \begin{cases} \alpha y X & X < \hat{D} \\ \hat{B} X^{\beta_2} + \frac{c}{r} & X \geq \hat{D}, \end{cases} \tag{A-30}$$

where $\hat{B}$ and β_2 are real-valued parameters. It is straightforward to show that

$$\beta_2 = \frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2} - \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2}\right)^2 + \frac{2r^n}{\tilde{\sigma}^2}} \tag{A-31}$$

$$\hat{B} = \left(\alpha y \hat{D} - \frac{c}{r^n}\right) \hat{D}^{-\beta_2}. \tag{A-32}$$

A.2.2. The valuation of tax benefits

The value of tax benefits $\hat{t}_i(X)$ corresponds to the value of debt with recovery rates equal to zero, and a coupon of $c\tau$ (and analogously for Case 2).

A.2.3. The valuation of default costs

As there are no continuous earnings associated with default costs, value function of default costs $\hat{b}_i(X)$ can be calculated as the value of a debt contract with recovery rates $1 - \alpha_B$ and $1 - \alpha_R$, respectively, and a coupon of zero. Case 2 can be treated analogously.

A.2.4. Firm value

Total firm value $\hat{f}_i$ in regime $i = B, R$ corresponds to the value of assets $y_i X$, plus the value of tax benefits from debt $\hat{t}_i(X)$, less the value of potential default costs $\hat{b}_i(X)$, i.e.,

$$\hat{f}_i(X) = X y_i + \hat{t}_i(X) - \hat{b}_i(X).$$

Analogously, for Case 2, we have

$$\hat{f}(X) = X y + \hat{t}(X) - \hat{b}(X).$$

A.2.5. The valuation of equity

The levered firm value equals the sum of debt and equity values. Hence, equity value $\hat{e}_i(X)$, $i = B, R$, may be written as

$$\hat{e}_i(X) = \hat{f}_i(X) - \hat{d}_i(X) = X y_i + \hat{t}_i(X) - \hat{b}_i(X) - \hat{d}_i(X), \tag{A-33}$$

or, for the Case 2,

$$\hat{e}(X) = \hat{f}(X) - \hat{d}(X) = X y + \hat{t}(X) - \hat{b}(X) - \hat{d}(X). \tag{A-34}$$

This is the closed-form expression for equity.

A.2.6. Default policy

Once debt has been issued, managers select the ex-post default policy that maximizes the value of equity. Formally, the default policy is determined by postulating that the derivative of the equity value has to be zero at the according default boundary. It is straightforward to calculate the first derivative of equity in closed form, using the derivative of the functional forms of the value of debt, default costs, and tax shield. The system to solve is for Case V1:

$$\begin{cases} \hat{e}'_B(\hat{D}^*_B) &= 0 \\ \hat{e}'_R(\hat{D}^*_R) &= 0. \end{cases} \tag{A-35}$$

We solve this problem numerically. For Case 2, the system is

$$\hat{e}'(D^*) = 0, \tag{A-36}$$

which is solvable in closed form.

Note that for a given coupon, all value functions can be calculated by following the approach up to system (A-35) or system (A-36), depending on the case. This fact will be used later for the calculation of the value of corporate securities of a firm consisting of both assets in place and an expansion option.

A.2.7. Capital structure

Denote by $\hat{f}^*_i(X)$ the firm value of a firm with only invested assets, given optimal ex-post default thresholds. The ex-ante optimal coupon of a firm solves
in Case V1

$$\hat{c}^* := \text{argmax}_{\hat{c}} \hat{f}^*_i(X), \tag{A-37}$$

and in Case 2

$$\hat{c}^* := \text{argmax}_{\hat{c}} \hat{f}^*(X). \tag{A-38}$$

A.3. The value of the growth option

Case G1: $X_R > X_B$**.** Recall that the system to solve is:

For $0 \leq X < X_B$:

$$\begin{cases} r^n_B G_B(X) &= \tilde{\mu}_B X G'_B(X) + \frac{1}{2}\tilde{\sigma}^2_B X^2 G''_B(X) + \tilde{\lambda}_B \left(G_R(X) - G_B(X)\right) \\ r^n_R G_R(X) &= \tilde{\mu}_R X G'_R(X) + \frac{1}{2}\tilde{\sigma}^2_R X^2 G''_R(X) + \tilde{\lambda}_R \left(G_B(X) - G_R(X)\right) \end{cases} \tag{A-39}$$

For $X_B \leq X < X_R$:

$$\begin{cases} G_B(X) &= sXy_B - K \\ r^n_R G_R(X) &= \tilde{\mu}_R X G'_R(X) + \frac{1}{2}\tilde{\sigma}^2_R X^2 G''_R(X) + \tilde{\lambda}_R \left(sXy_B - K - G_R(X)\right) \end{cases} \tag{A-40}$$

For $X \geq X_R$:

$$\begin{cases} G_B(X) &= sXy_B - K \\ G_R(X) &= sXy_R - K, \end{cases} \tag{A-41}$$

subject to the boundary conditions:

$$\lim_{X \searrow 0} G_i(X) = 0, \quad i = B, R \tag{A-42}$$

$$\lim_{X \searrow X_B} G_R(X) = \lim_{X \nearrow X_B} G_R(X) \tag{A-43}$$

$$\lim_{X \searrow X_B} G'_R(X) = \lim_{X \nearrow X_B} G'_R(X) \tag{A-44}$$

$$\lim_{X \nearrow X_R} G_R(X) = sX_R y_R - K \tag{A-45}$$

$$\lim_{X \nearrow X_B} G_B(X) = sX_B y_B - K \tag{A-46}$$

The functional form of the solution is given by

$$G_i(X) = \begin{cases} \bar{A}_{i3} X^{\gamma_3} + \bar{A}_{i4} X^{\gamma_4} & X < X_B, & i = B, R \\ \bar{C}_1 X^{\beta_1^R} + \bar{C}_2 X^{\beta_2^R} + \bar{C}_3 X + \bar{C}_4 & X_B \leq X < X_R, & i = R \\ sXy_i - K & X \geq X_i & i = B, R, \end{cases} \tag{A-47}$$

where $\bar{A}_{B3}, \bar{A}_{B4}, \bar{A}_{R1}, \bar{A}_{R2}, \bar{C}_1, \bar{C}_2, \bar{C}_3, \bar{C}_4, \gamma_3, \gamma_4, \beta_1^R$, and β_2^R are real-valued parameters to be determined. The notation$^-$ indicates that a parameter refers to the value of the growth option (and only to the value of the growth option). We first consider the region $X < X_B$, and use the standard approach of plugging in the functional form $G_i(X) = \bar{A}_{i3} X^{\gamma_3} + \bar{A}_{i4} X^{\gamma_4}$ into both equations of (A-39). Comparison of coefficients yields that $\bar{A}_{Rk}$ is always a multiple of $\bar{A}_{Bk}$, $k = 3, 4$, with the factor $\bar{l}_k := \frac{1}{\tilde{\lambda}_B}(r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B \gamma_k - \frac{1}{2}\tilde{\sigma}_B^2 \gamma_k(\gamma_k - 1))$, i.e., $\bar{A}_{Bk} = \bar{l}_k \bar{A}_{Rk}$. Note that even though the factor $\bar{l}_k$ is of similar structure as the one found in the calculation of the value of debt of a firm with only invested assets, their values differ due to the different roots γ_i in the formulae. Using this relationship and comparing coefficients again, we find that γ_3 and γ_4 correspond to the positive roots of the quartic equation

$$\left(\tilde{\mu}_R \gamma + \frac{1}{2}\tilde{\sigma}_R^2 \gamma(\gamma - 1) - \tilde{\lambda}_R - r_R^n\right)\left(\tilde{\mu}_B \gamma + \frac{1}{2}\tilde{\sigma}_B^2 \gamma(\gamma - 1) - \tilde{\lambda}_B - r_B^n\right) = \tilde{\lambda}_R \tilde{\lambda}_B. \tag{A-48}$$

The reason for taking the positive roots being that the option value has to approach zero as the earnings approaches zero.

Next, we consider the region $X_B \leq X < X_R$. Note that in the case of interest the Markov chain is in recession (otherwise, the solution is already known). Again, plugging in the functional form $G_R(X) = \bar{C}_1 X^{\beta_1} + \bar{C}_2 X^{\beta_2} + \bar{C}_3 X + \bar{C}_4$ into the second equation of (A-40), we find by comparison of coefficients that

$$\begin{aligned} \beta_{1,2}^R &= \frac{1}{2} - \frac{\tilde{\mu}_R}{\tilde{\sigma}_R^2} \pm \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}_R}{\tilde{\sigma}_R^2}\right)^2 + \frac{2\left(r_R^n + \tilde{\lambda}_R\right)}{\tilde{\sigma}_R^2}} \\ \bar{C}_3 &= \frac{s\tilde{\lambda}_R y_B}{r_R^n - \tilde{\mu}_R + \tilde{\lambda}_R} \\ \bar{C}_4 &= -\frac{K\tilde{\lambda}_R}{r_R^n + \tilde{\lambda}_R}. \end{aligned} \tag{A-49}$$

It is left to solve for the unknown parameters $\bar{A}_{B3}, \bar{A}_{B4}, \bar{C}_1$ and $\bar{C}_2$. Plugging in the functional form (A-47) into conditions (A-43)-(A-46) yields

$$\begin{aligned}
\bar{C}_1 X_B^{\beta_1^R} + \bar{C}_2 X_B^{\beta_2^R} + \bar{C}_3 X_B + \bar{C}_4 &= \bar{l}_1 \bar{A}_{B3} X_B^{\gamma_3} + \bar{l}_2 \bar{A}_{B4} X_B^{\gamma_4} && \text{(A-50)}\\
\bar{C}_1 \beta_1^R X_B^{\beta_1^R} + \bar{C}_2 \beta_2^R X_B^{\beta_2^R} + \bar{C}_3 X_B &= \bar{l}_1 \bar{A}_{B3} \gamma_3 X_B^{\gamma_3} + \bar{l}_2 \gamma_4 \bar{A}_{B4} X_B^{\gamma_4} && \text{(A-51)}\\
\bar{C}_1 X_R^{\beta_1^R} + \bar{C}_2 X_R^{\beta_2^R} + \bar{C}_3 X_R + \bar{C}_4 &= s y_R X_R - K && \text{(A-52)}\\
\bar{A}_{B3} X_B^{\gamma_3} + \bar{A}_{B4} X_B^{\gamma_4} &= s y_B X_B - K && \text{(A-53)}
\end{aligned}$$

This four-dimensional system is linear in its four unknowns $\bar{A}_{B3}, \bar{A}_{B4}, \bar{C}_1$ and $\bar{C}_2$. We define the matrices

$$\begin{aligned}
\bar{M} &:= \begin{bmatrix} \bar{l}_1 X_B^{\gamma_3} & \bar{l}_2 X_B^{\gamma_4} & -X_B^{\beta_1^R} & -X_B^{\beta_2^R} \\ \bar{l}_1 \gamma_3 X_B^{\gamma_3} & \bar{l}_2 \gamma_4 X_B^{\gamma_4} & -\beta_1^R X_B^{\beta_1^R} & -\beta_2^R X_B^{\beta_2^R} \\ 0 & 0 & X_R^{\beta_1^R} & X_R^{\beta_2^R} \\ X_B^{\gamma_3} & X_B^{\gamma_4} & 0 & 0 \end{bmatrix} \\
\bar{b} &:= \begin{bmatrix} \bar{C}_3 X_B + \bar{C}_4 \\ \bar{C}_3 X_B \\ -\bar{C}_3 X_R - \bar{C}_4 + s y_R X_R - K \\ s y_B X_B - K \end{bmatrix},
\end{aligned}$$

such that $\bar{M} \begin{bmatrix} \bar{A}_{B3} & \bar{A}_{B4} & \bar{C}_1 & \bar{C}_2 \end{bmatrix}^T = \bar{b}$. Hence the solution to the remaining unknowns is given by

$$\begin{bmatrix} \bar{A}_{B3} & \bar{A}_{B4} & \bar{C}_1 & \bar{C}_2 \end{bmatrix}^T = \bar{M}^{-1} \bar{b}. \tag{A-54}$$

Note that the relative price change sensitivity is

$$\frac{G_i'(X)}{G_i(X)} = \begin{cases} \frac{\gamma_3 \bar{A}_{i3} X^{\gamma_3 - 1} + \bar{A}_{i4} \gamma_4 X^{\gamma_4 - 1}}{\bar{A}_{i3} X^{\gamma_3} + \bar{A}_{i4} X^{\gamma_4}} & X < X_B, \quad i = B, R \\ \frac{\bar{C}_1 \beta_1 X^{\beta_1 - 1} + \bar{C}_2 \beta_2 X^{\beta_2 - 1} + \bar{C}_3}{\bar{C}_1 X^{\beta_1} + \bar{C}_2 X^{\beta_2} + \bar{C}_3 X + \bar{C}_4} & X_B \le X < X_R, \quad i = R \\ \frac{s y_i}{s y_i X - K} & X \ge X_i \quad i = B, R. \end{cases} \tag{A-55}$$

Finally, consider the unlevered value of the growth option, whose optimal exercise boundaries are determined by the additional boundary conditions (23)-(24):

$$\begin{aligned}
\lim_{X \nearrow X_R^{unlev}} G_R'(X) &= s y_R && \text{(A-56)}\\
\lim_{X \nearrow X_B^{unlev}} G_B'(X) &= s y_B. && \text{(A-57)}
\end{aligned}$$

The calculations are the same up to system (A-49). System (A-50)-(A-53) is augmented by the two equations corresponding to the additional boundary conditions:

$$\begin{aligned}
\bar{C}_1^{unlev} \beta_1^R \left(X_R^{unlev}\right)^{\beta_1^R - 1} + \bar{C}_2^{unlev} \beta_2^R \left(X_R^{unlev}\right)^{\beta_2^R - 1} + \bar{C}_3 &= s y_R && \text{(A-58)}\\
\bar{A}_{B3}^{unlev} \gamma_3 \left(X_B^{unlev}\right)^{\gamma_3 - 1} + \bar{A}_{B4}^{unlev} \gamma_4 \left(X_B^{unlev}\right)^{\gamma_4 - 1} &= s y_B. && \text{(A-59)}
\end{aligned}$$

The full system is six-dimensional with the six unknowns $\bar{A}_{B3}^{unlev}$, $\bar{A}_{B4}^{unlev}$, $\bar{C}_1^{unlev}$, $\bar{C}_2^{unlev}$ X_B^{unlev}, and X_R^{unlev}, linear in the first four unknowns, and non-linear in the last two unknowns. It is solved numerically, using relation (A-54) for any given pair of exercise boundaries in the numerical solution algorithm.

Case 2: Only 1 regime. Denote the investment boundary by X_1. We find that the system to solve is given by:

$$\begin{array}{rcll} rG(X) &=& \tilde{\mu}XG'(X) + \frac{\tilde{\sigma}^2}{2}X^2G''(X) & X < X_1 \\ G(X) &=& sXy - K & X \geq X_1 \end{array} \tag{A-60}$$

The boundary conditions are given by a value matching condition and the fact that he option must become worthless as the asset value approaches zero:

$$\lim_{X \nearrow X_1} G(X) = syX_1 - K \tag{A-61}$$

$$\lim_{X \searrow 0} G(X) = 0 \tag{A-62}$$

The functional form of the solution is

$$G(X) = \begin{cases} \bar{A}X^{\beta_1} & X < X_1 \\ sXy - K & X \geq X_1, \end{cases} \tag{A-63}$$

where $\bar{A}$ and β_1 are real-valued parameters to be determined. It is then straightforward to show that

$$\beta_1 = \frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2} + \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2}\right)^2 + \frac{2r}{\tilde{\sigma}^2}} \tag{A-64}$$

$$\bar{A} = (syX_1 - K)\, X_1^{-\beta_1}, \tag{A-65}$$

which is the solution for the option. The relative price change sensitivity of the option is

$$\frac{G'(X)}{G(X)} = \begin{cases} \frac{\beta_1}{X} & X < X_1 \\ \frac{sy}{syX-K} & X \geq X_1. \end{cases} \tag{A-66}$$

A.4. Firms with invested assets and expansion options

A.4.1. The valuation of corporate debt

Case A1: $D_B < D_R$, $\hat{D}_B < \hat{D}_R$, and $X_R > X_B$. This case constitutes the one presented in the main text. For brevity of notation, define $\bar{s} := s + 1$. Recall that the system to solve is:
For $0 \leq X \leq D_B$:

$$\begin{cases} d_B(X) = \alpha_B\left(Xy_B + G_B^{unlev}(X)\right) \\ d_R(X) = \alpha_R\left(Xy_R + G_R^{unlev}(X)\right) \end{cases} \tag{A-67}$$

For $D_B < X \leq D_R$:

$$\begin{cases} r_B^n d_B(X) = c + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) + \tilde{\lambda}_B\left(\alpha_R\left(Xy_R + G_R^{unlev}(X)\right) - d_B(X)\right) \\ d_R(X) = \alpha_R\left(Xy_R + G_R^{unlev}(X)\right) \end{cases} \tag{A-68}$$

For $D_R < X < X_B$:

$$\left\{\begin{array}{lcl} r^n_B d_B\left(X\right) & = & c+\tilde{\mu}_B X d'_B\left(X\right)+\frac{1}{2}\tilde{\sigma}^2_B X^2 d''_B\left(X\right)+\tilde{\lambda}_B\left(d_R\left(X\right)-d_B\left(X\right)\right) \\ r^n_R d_R\left(X\right) & = & c+\tilde{\mu}_R X d'_R\left(X\right)+\frac{1}{2}\tilde{\sigma}^2_R X^2 d''_R\left(X\right)+\tilde{\lambda}_R\left(d_B\left(X\right)-d_R\left(X\right)\right) \end{array}\right. \tag{A-69}$$

For $X_B \leq X < X_R$:

$$\left\{\begin{array}{lcl} d_B\left(X\right) & = & \hat{d}_B\left(\bar{s}X-\frac{K}{y_B}\right) \\ r d_R\left(X\right) & = & c+\tilde{\mu}_R X d'_R\left(X\right)+\frac{1}{2}\tilde{\sigma}^2_R X^2 d''_R\left(X\right)+\tilde{\lambda}_R\left(\hat{d}_B\left(\bar{s}X-\frac{K}{y_B}\right)-d_R\left(X\right)\right) \end{array}\right. \tag{A-70}$$

For $X \geq X_R$:

$$\left\{\begin{array}{lcl} d_B\left(X\right) & = & \hat{d}_B\left(\bar{s}X-\frac{K}{y_B}\right) \\ d_R\left(X\right) & = & \hat{d}_R\left(\bar{s}X-\frac{K}{y_R}\right). \end{array}\right. \tag{A-71}$$

The system is subject to the following boundary conditions:

$$\begin{array}{rcl} \lim\limits_{X\searrow D_R} d_B\left(X\right) & = & \lim\limits_{X\nearrow D_R} d_B\left(X\right) \\ \lim\limits_{X\searrow D_R} d'_B\left(X\right) & = & \lim\limits_{X\nearrow D_R} d'_B\left(X\right) \\ \lim\limits_{X\searrow D_B} d_B\left(X\right) & = & \alpha_B\left(D_B y_B+G^{unlev}_B\left(D_B\right)\right) \\ \lim\limits_{X\searrow D_R} d_R\left(X\right) & = & \alpha_R\left(D_R y_R+G^{unlev}_R\left(D_R\right)\right) \\ \lim\limits_{X\searrow X_B} d_R\left(X\right) & = & \lim\limits_{X\nearrow X_B} d_R\left(X\right) \\ \lim\limits_{X\searrow X_B} d'_R\left(X\right) & = & \lim\limits_{X\nearrow X_B} d'_R\left(X\right) \\ \lim\limits_{X\nearrow X_B} d_B\left(X\right) & = & \hat{d}_B\left(\bar{s}X_B-\dfrac{K}{y_B}\right) \\ \lim\limits_{X\nearrow X_R} d_R\left(X_R\right) & = & \hat{d}_R\left(\bar{s}X_R-\dfrac{K}{y_R}\right). \end{array} \tag{A-72}$$

In order to solve this system, we start with the functional form of the solution:

$$d_i\left(X\right)=\left\{\begin{array}{lll} \alpha_i\left(Xy_i+G^{unlev}_i\left(X\right)\right) & X\leq D_i & i=B,R, \\ C_1X^{\beta^B_1}+C_2X^{\beta^B_2}+C_3X+C_4 & & \\ \quad +C_5X^{\gamma_3}+C_6X^{\gamma_4} & D_B<X\leq D_R, & i=B \\ A_{i1}X^{\gamma_1}+A_{i2}X^{\gamma_2} & & \\ \quad +A_{i3}X^{\gamma_3}+A_{i4}X^{\gamma_4}+A_5 & D_R<X\leq X_B, & i=B,R \\ B_1X^{\beta^R_1}+B_2X^{\beta^R_2}+Z\left(X\right) & X_B<X\leq X_R, & i=R \\ \hat{d}_i\left(\bar{s}X-\frac{K}{y_i}\right) & X>X_i, & i=B,R, \end{array}\right. \tag{A-73}$$

where $A_{B1}, A_{B2}, A_{R1}, A_{R2}, C_1, C_2, C_3, C_4, C_5, C_6, B_1, B_2, \beta^B_1, \beta^B_2, \beta^R_1, \beta^R_2, \gamma_3$, and γ_4 are real-valued parameters to be determined (or to be confirmed). The function $Z(X)$ as stated in the sixth line of (A-73) is of closed form. It will be given explicitly in the following calculations.

We first consider the region $X_B \geq X > D_R$. Using the standard approach of plugging in the functional form $d_i(X) = A_{i1}X^{\gamma_1} + A_{i2}X^{\gamma_2} + A_{i3}X^{\gamma_3} + A_{i4}X^{\gamma_4} + A_{i5}$ into both equations of (A-69), comparing coefficients, and solving, we confirm that

$$A_{i5} = \frac{c\left(r_j^n + \tilde{\lambda}_i + \tilde{\lambda}_j\right)}{r_i^n r_j^n + r_j^n\tilde{\lambda}_i + r_i^n\tilde{\lambda}_j}, \tag{A-74}$$

and we find again that A_{Rk} is always a multiple of A_{Bk}, $k = 1, \ldots, 4$, with the factor $l_k := \frac{1}{\tilde{\lambda}_B}(r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B\gamma_k - \frac{1}{2}\tilde{\sigma}_B^2\gamma_k(\gamma_k - 1))$, i.e., $A_{Bk} = l_k A_{Rk}$. Using this relationship and comparing coefficients, we find that $\gamma_1, \gamma_2, \gamma_3$, and γ_4 correspond to the roots of the quartic equation (A-19), which is:

$$\left(\tilde{\mu}_R\gamma + \frac{1}{2}\tilde{\sigma}_R^2\gamma(\gamma-1) - \tilde{\lambda}_R - r_R^n\right)\left(\tilde{\mu}_B\gamma + \frac{1}{2}\tilde{\sigma}_B^2\gamma(\gamma-1) - \tilde{\lambda}_B - r_B^n\right) = \tilde{\lambda}_R\tilde{\lambda}_B. \tag{A-75}$$

Recall that by arguments of Guo (2001), we know that this quartic equation always has four distinct real roots, two of them being negative, and two positive. The value of debt in both regimes will be subject to boundary conditions from both below (default) and above (exercise of expansion option). In order to meet all boundary conditions, we need four terms with the according factors A_{ik} as well as exponents γ_k, which requires usage of all four roots of (A-75). The no-bubbles condition is already implemented in the value function of a firm with only invested assets $\hat{d}_i$, and, hence, does not need to be imposed again. The unknown parameters left for this region are $A_{Bk},\ k = 1, \ldots, 4$.

Next, we consider the region $D_B \leq X \leq D_R$, i.e., the realized state of the Markov chain is boom (in recession, the solution is already known by the second equation of system (A-68)). Plugging in the functional form $d_B(X) = C_1X^{\beta_1^B} + C_2X^{\beta_2^B} + C_3X + C_4 + C_5X^{\gamma_3} + C_6X^{\gamma_4}$ into the second equation of (A-68), we find by comparison of coefficients that

$$\begin{aligned}
\beta_{1,2}^B &= \frac{1}{2} - \frac{\tilde{\mu}_B}{\tilde{\sigma}_B^2} \pm \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}_B}{\tilde{\sigma}_B^2}\right)^2 + \frac{2\left(r_B^n + \tilde{\lambda}_B\right)}{\tilde{\sigma}_B^2}} && \text{(A-76)}\\
C_3 &= \frac{\tilde{\lambda}_B\alpha_R y_R}{r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B} && \text{(A-77)}\\
C_4 &= \frac{c}{r_B^n + \tilde{\lambda}_B} && \text{(A-78)}\\
C_5 &= \frac{\tilde{\lambda}_B\alpha_R\bar{l}_1\bar{A}_{B3}^{unlev}}{r_B^n - \tilde{\mu}_B\gamma_3 - \frac{1}{2}\tilde{\sigma}_B^2\gamma_3\left(\gamma_3 - 1\right) + \tilde{\lambda}_B} && \text{(A-79)}\\
C_6 &= \frac{\tilde{\lambda}_B\alpha_R\bar{l}_2\bar{A}_{B4}^{unlev}}{r_B^n - \tilde{\mu}_B\gamma_4 - \frac{1}{2}\tilde{\sigma}_B^2\gamma_4\left(\gamma_4 - 1\right) + \tilde{\lambda}_B}. && \text{(A-80)}
\end{aligned}$$

We require again that $r_B^n + \tilde{\lambda}_B - \tilde{\mu}_B \neq 0$. Note that the denominators of C_5 and C_6 are different from zero as long as the Markov chain I is recurrent, i.e., if $\tilde{\lambda}_i > 0$, $i = B, R$ (see equation (A-75)). The parameters $\beta_{1,2}^B, C_3$, and C_4 are the same as for a firm with only invested assets, cf. Appendix A.2, equations (A-20). C_5 and C_6 are influenced by the parameters in the solution of the growth option, $\bar{l}_1, \bar{l}_2, \bar{A}_{B3}^{unlev}$, and $\bar{A}_{B4}^{unlev}$ (see Appendix A.3). The two additional terms of the solution for this region, $C_5X^{\gamma_3}$ and $C_6X^{\gamma_4}$, reflect the fact that the firm does not only consist of assets in place, but also of the growth option. As debtholders get also a fraction of the growth option's value at regime-switching induced default, the value of the option directly influences the solution in this region. This influence explains the occurrence of the growth option parameters in C_5 and C_6, as well as the use of the same exponents as in the calculation of the value of the option, γ_3 and γ_4. Note that the approach and the intuition regarding the exponents γ_3 and γ_4 for this region are completely different than for the previously discussed region $X_B \geq X > D_R$, where these exponents occur only due to the valuation of debt itself, independent of the growth option, and must be calculated as a part of the solution. The unknown parameters left for this region are C_1 and C_2.

Finally, consider the region $X_B < X \leq X_R$ for $i = R$. The corresponding differential equation is (see (A-70)):

$$r_R^n d_R(X) = c + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) + \tilde{\lambda}_R \hat{d}_B(\bar{s}X - \frac{K}{y_B}). \tag{A-81}$$

In order to solve this inhomogeneous differential equation, we use a standard approach by first finding a fundamental system of solutions of the homogenous differential equation, and then calculating the solution of the inhomogeneous equation as the sum of the solutions of the homogenous equation and a particular solution of the nonhomogeneous equation. A reference for this approach is Polyanin and Zaitsev (2003), pages 21-23.[31]

(A-81) is equivalent to

$$X^2 d_R''(X) + \frac{2\tilde{\mu}_R}{\tilde{\sigma}_R^2} X d_R'(X) - \frac{2\left(r_R^n + \tilde{\lambda}_R\right)}{\tilde{\sigma}_R^2} d_R(X) = -\frac{2c}{\tilde{\sigma}_R^2} - \frac{2\tilde{\lambda}_R}{\tilde{\sigma}_R^2}\hat{d}_B(\bar{s}X - \frac{K}{y_B}). \tag{A-82}$$

Therefore, the according homogenous differential equation is

$$X^2 d_R''(X) + \frac{2\tilde{\mu}_R}{\tilde{\sigma}_R^2} X d_R'(X) - \frac{2\left(r_R^n + \tilde{\lambda}_R\right)}{\tilde{\sigma}_R^2} d_R(X) = 0. \tag{A-83}$$

A fundamental system of solutions is given by $\{z_1, z_2\}$, with

$$\begin{aligned} z_1 &:= X^{\beta_1^R}, \\ z_2 &:= X^{\beta_2^R}, \end{aligned}$$

and

$$\beta_{1,2}^R = \frac{1}{2} - \frac{\tilde{\mu}_R}{\tilde{\sigma}_R^2} \pm \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}_R}{\tilde{\sigma}_R^2}\right)^2 + \frac{2\left(r_R^n + \tilde{\lambda}_R\right)}{\tilde{\sigma}_R^2}}. \tag{A-84}$$

These solutions can be calculated by plugging the functional form into the homogenous ODE (A-83), and solving for $\beta_{1,2}^R$.

[31]Technically, the above explained two-step procedure for the solution of the ODE is required due to the assumption that the exercise of the option is financed by selling a part of the assets in place, resulting in the fact that the function $\hat{d}$ in the ODE is not evaluated at X, but at $\bar{s}X - \frac{K}{y_B}$. Under the alternative assumption that the exercise of the option is equity-financed, the function $\hat{d}$ is evaluated at a multiple of X instead. In this case, we can exploit the additive nature of the ODE, and calculate the solution as a weighted sum of solutions, including the value function of debt of a firm with only invested assets. The functional form for the region $X_B < X \leq X_R$ is then comparable to the one for the region $D_B \leq X \leq D_R$.

For notational convenience, we now define $f_2 := X^2$, $f_1 := \frac{2\tilde{\mu}_R}{\tilde{\sigma}_R^2}X$, $f_0 := -\frac{2\left(r_R^n+\tilde{\lambda}_R\right)}{\tilde{\sigma}_R^2}$, and

$$g\left(X\right) := -\frac{2c}{\tilde{\sigma}_R^2} - \frac{2\tilde{\lambda}_R}{\tilde{\sigma}_R^2}\hat{d}_B(\bar{s}X - \frac{K}{y_B}). \tag{A-85}$$

These notations allow to write the ODE (A-82) as:

$$f_2 d_R''(X) + f_1 d_R'(X) + f_0 d_R(X) = g(X). \tag{A-86}$$

The general solution of this inhomogeneous ODE is given by

$$d_R\left(X\right) = B_1 z_1 + B_2 z_2 + z_2 \underbrace{\int z_1 \frac{g}{f_2}\frac{\mathrm{d}X}{W}}_{=:I_1(X)} - z_1 \underbrace{\int z_2 \frac{g}{f_2}\frac{\mathrm{d}X}{W}}_{=:I_2(X)}, \tag{A-87}$$

where $W = z_1 z_2' - z_2 z_1'$ is the Wronskian determinant, and B_1 and B_2 are coefficients (see e.g. Polyanin and Zaitsev (2003), page 22, (7)). The first two terms are a linear combination of the solutions of the homogenous ODE, and the last two terms are a particular solution of the inhomogeneous ODE.

We start by calculating the Wronskian determinant

$$\begin{aligned} W &= z_1 z_2' - z_2 z_1' \\ &= \beta_2^R X^{\beta_1^R} X^{\beta_2^R - 1} - \beta_1^R X^{\beta_1^R - 1} X^{\beta_2^R} \\ &= \left(\beta_2^R - \beta_1^R\right) X^{\beta_1^R + \beta_2^R - 1}. \end{aligned} \tag{A-88}$$

The integral $I_1(X)$ is, hence:

$$\begin{aligned}
I_1(X) &= \int z_1 \frac{g}{f_2}\frac{\mathrm{d}X}{W} \\
&= \int X^{\beta_1^R} X^{-2}\frac{1}{\beta_2^R-\beta_1^R} X^{1-\beta_1^R-\beta_2^R} g(X)\mathrm{d}X \\
&= \frac{1}{\beta_2^R-\beta_1^R}\int x^{-1-\beta_2^R} g(X)\mathrm{d}X && \text{(A-89)} \\
&= \frac{1}{\beta_2^R-\beta_1^R}\int x^{-1-\beta_2^R} \quad -\frac{2c}{\tilde{\sigma}_R^2} - \frac{2\tilde{\lambda}_R}{\tilde{\sigma}_R^2}\hat{d}_B\left(\bar{s}X-\frac{K}{y_B}\right)\Bigg)\mathrm{d}X \\
&= \frac{1}{\beta_2^R-\beta_1^R}\int x^{-1-\beta_2^R}\Bigg(-\frac{2c}{\tilde{\sigma}_R^2} \\
&\quad -\frac{2\tilde{\lambda}_R}{\tilde{\sigma}_R^2}\left\{\hat{A}_{B1}\left(\bar{s}X-\frac{K}{y_B}\right)^{\gamma_1}+\hat{A}_{B2}\left(\bar{s}X-\frac{K}{y_B}\right)^{\gamma_2}+\frac{c}{r}\right\}\Bigg)\mathrm{d}X \\
&= -\frac{2\tilde{\lambda}_R\hat{A}_{B1}}{\left(\beta_2^R-\beta_1^R\right)\tilde{\sigma}_R^2}\underbrace{\int X^{-1-\beta_2^R}\left(\bar{s}X-\frac{K}{y_B}\right)^{\gamma_1}\mathrm{d}X}_{=:I_{11}(X)} \\
&\quad -\frac{2\tilde{\lambda}_R\hat{A}_{B2}}{\left(\beta_2^R-\beta_1^R\right)\tilde{\sigma}_R^2}\underbrace{\int X^{-1-\beta_2^R}\left(\bar{s}X-\frac{K}{y_B}\right)^{\gamma_2}\mathrm{d}X}_{=:I_{12}(X)} && \text{(A-90)} \\
&\quad +\frac{2c\left(\tilde{\lambda}_R+r_R^n\right)}{\left(\beta_2^R-\beta_1^R\right)r_R^n\beta_2^R\tilde{\sigma}_R^2}X^{-\beta_2^R}.
\end{aligned}$$

We use the definition of the function $g(X)$, see (A-85), and the solution of the debt value of a firm with only invested assets $\hat{d}_R(\cdot)$, see Appendix A.2, (A-17).

The integrals $I_{11}(X)$ and $I_{12}(X)$ can be evaluated immediately with standard computer algebra packages. Alternatively, using the integral representation of Gauss' hypergeometric function ${}_2F_1(\cdot,\cdot,\cdot;\cdot)$, we can write the closed-form solution of the integrals as

$$\begin{aligned}
I_{11}(X) &= \frac{1}{\gamma_1-\beta_2^R}\bar{s}^{\gamma_1}X^{\gamma_1-\beta_2^R}{}_2F_1\left(-\gamma_1,\beta_2^R-\gamma_1,\beta_2^R-\gamma_1+1;-\frac{K}{\bar{s}Xy_B}\right), && \text{(A-91)} \\
I_{12}(X) &= \frac{1}{\gamma_2-\beta_2^R}\bar{s}^{\gamma_2}X^{\gamma_2-\beta_2^R}{}_2F_1\left(-\gamma_2,\beta_2^R-\gamma_2,\beta_2^R-\gamma_2+1;-\frac{K}{\bar{s}Xy_B}\right). && \text{(A-92)}
\end{aligned}$$

Plugging the solutions (A-91) and (A-92) back into the expression for the integral I_1, (A-90) yields

$$\begin{aligned}
I_1(X) &= -\frac{2\tilde{\lambda}_R\hat{A}_{B1}}{\left(\beta_2^R-\beta_1^R\right)\tilde{\sigma}_R^2}\frac{1}{\gamma_1-\beta_2^R}\bar{s}^{\gamma_1}X^{\gamma_1-\beta_2^R}{}_2F_1\left(-\gamma_1,\beta_2^R-\gamma_1,\beta_2^R-\gamma_1+1;-\frac{K}{\bar{s}Xy_B}\right) \\
&\quad -\frac{2\tilde{\lambda}_R\hat{A}_{B2}}{\left(\beta_2^R-\beta_1^R\right)\tilde{\sigma}_R^2}\frac{1}{\gamma_2-\beta_2^R}\bar{s}^{\gamma_2}X^{\gamma_2-\beta_2^R}{}_2F_1\left(-\gamma_2,\beta_2^R-\gamma_2,\beta_2^R-\gamma_2+1;-\frac{K}{\bar{s}Xy_B}\right) && \text{(A-93)} \\
&\quad +\frac{2c\left(\tilde{\lambda}_R+r\right)}{\left(\beta_2^R-\beta_1^R\right)r\beta_2^R\tilde{\sigma}_R^2}X^{-\beta_2^R}.
\end{aligned}$$

Similarly, we find for the second integral $I_2(X)$:

$$\begin{aligned}
I_2(X) &= -\frac{2\tilde{\lambda}_R \hat{A}_{B1}}{\left(\beta_2^R-\beta_1^R\right)\tilde{\sigma}_R^2}\frac{1}{\gamma_1-\beta_1^R}\bar{s}^{\gamma_1}X^{\gamma_1-\beta_1^R}{}_2F_1\left(-\gamma_1,\beta_1^R-\gamma_1,\beta_2^R-\gamma_1+1;-\frac{K}{\bar{s}Xy_B}\right)\\
&\quad -\frac{2\tilde{\lambda}_R \hat{A}_{B2}}{\left(\beta_2^R-\beta_1^R\right)\tilde{\sigma}_R^2}\frac{1}{\gamma_2-\beta_1^R}\bar{s}^{\gamma_2}X^{\gamma_2-\beta_1^R}{}_2F_1\left(-\gamma_2,\beta_1^R-\gamma_2,\beta_2^R-\gamma_2+1;-\frac{K}{\bar{s}Xy_B}\right)\\
&\quad +\frac{2c\left(\tilde{\lambda}_R+r\right)}{\left(\beta_2^R-\beta_1^R\right)r\beta_1^R\tilde{\sigma}_R^2}X^{-\beta_1^R}.
\end{aligned} \tag{A-94}$$

Plugging (A-93) and (A-94) into (A-87), we finally obtain the solution

$$d_R(X) = B_1X^{\beta_1^R} + B_2X^{\beta_2^R} + Z(X), \tag{A-95}$$

with

$$\begin{aligned}
Z(X) &= \frac{2}{\beta_1\beta_2\tilde{\sigma}_R^2}\frac{c}{r}\left(\tilde{\lambda}_R+r\right)\\
&\quad +\sum_{i,k=1,2}\frac{2(-1)^{i+1}\bar{s}^{\gamma_k}\hat{A}_{Bk}}{\tilde{\sigma}_R^2\left(\beta_2^R-\beta_1^R\right)\left(\gamma_k-\beta_i^R\right)}X^{\gamma_k}{}_2F_1\left(-\gamma_k,\beta_i^R,\beta_i^R-\gamma_k+1;-\frac{K}{\bar{s}Xy_B}\right),
\end{aligned} \tag{A-96}$$

for some parameters B_1 and B_2 determined by the boundary conditions.

In order to treat the boundary conditions, we also need the first derivative of Z:

$$\begin{aligned}
Z'(X) &= \frac{\mathrm{d}}{\mathrm{d}X}Z(X)\\
&= \frac{\mathrm{d}}{\mathrm{d}X}\left(X^{\beta_2^R}I_1(X) - X^{\beta_2^R}I_2(X)\right)\\
&= \beta_2^R X^{\beta_2^R}I_1(X) + \frac{1}{\beta_2^R-\beta_1^R}X^{\beta_2^R}X^{-1-\beta_1^R}g(X)\\
&\quad -\beta_1^R X^{\beta_1^R}I_2(X) - \frac{1}{\beta_2^R-\beta_1^R}X^{\beta_1^R}X^{-1-\beta_1^R}g(X)\\
&= \beta_2^R X^{\beta_2^R}I_1(X) - \beta_1^R X^{\beta_1^R}I_2(X)\\
&= \sum_{i,k=1,2}\frac{2(-1)^{i+1}\bar{s}^{\gamma_k}\hat{A}_{Bk}\beta_i^R}{\tilde{\sigma}_R^2\left(\beta_2^R-\beta_1^R\right)\left(\gamma_k-\beta_i^R\right)}X^{\gamma_k}{}_2F_1\left(-\gamma_k,\beta_i^R,\beta_i^R-\gamma_k+1;-\frac{K}{\bar{s}Xy_B}\right).
\end{aligned} \tag{A-97}$$

To solve for the unknown parameters $A_{B1}, A_{B2}, A_{B3}, A_{B4}, C_1, C_2, B_1$ and B_2, we plug the functional form (A-73) into the system of boundary conditions (A-72):

$$\begin{aligned}
\sum_{k=1}^{4} A_{Bk} D_R^{\gamma_k} + A_5 &= C_1 D_R^{\beta_1^B} + C_2 D_R^{\beta_2^B} + C_3 X + C_4 + C_5 X^{\gamma_3} + C_6 X^{\gamma_4} \\
\sum_{k=1}^{4} A_{Bk} \gamma_k D_R^{\gamma_k} &= C_1 \beta_1^B D_R^{\beta_1^B} + C_2 \beta_2^B D_R^{\beta_2^B} + C_3 X + C_5 \gamma_3 X^{\gamma_3} + C_6 \gamma_4 X^{\gamma_4} \\
\alpha_B \left(D_B y_B + G_B^{unlev}(D_B) \right) &= C_1 D_B^{\beta_1^B} + C_2 D_B^{\beta_2^B} + C_3 D_B + C_4 + C_5 D_B^{\gamma_3} + C_6 D_B^{\gamma_4} \\
\sum_{k=1}^{4} l_k A_{Bk} D_R^{\gamma_k} + A_5 &= \alpha_R \left(D_R y_R + G_R^{unlev}(D_R) \right) \\
\sum_{k=1}^{4} l_k A_{Bk} X_B^{\gamma_k} + A_5 &= B_1 X_B^{\beta_1^R} + B_2 X_B^{\beta_2^R} + Z(X_B) \\
\sum_{k=1}^{4} l_k A_{Bk} \gamma_k X_B^{\gamma_k} &= B_1 \beta_1^R X_B^{\beta_1^R} + B_2 \beta_2^R X_B^{\beta_2^R} + X_B Z'(X_B) \\
\sum_{k=1}^{4} A_{Bk} X_B^{\gamma_k} + A_5 &= \hat{d}_B \left(\bar{s} X_B - \frac{K}{y_B} \right) \\
B_1 X_R^{\beta_1^R} + B_2 X_R^{\beta_2^R} + Z(X_R) &= \hat{d}_R \left(\bar{s} X_R - \frac{K}{y_R} \right).
\end{aligned} \tag{A-98}$$

Using matrix notation, we can write

$$M := \begin{bmatrix}
D_R^{\gamma_1} & D_R^{\gamma_2} & D_R^{\gamma_3} & D_R^{\gamma_4} & -D_R^{\beta_1^B} & -D_R^{\beta_2^B} & 0 & 0 \\
\gamma_1 D_R^{\gamma_1} & \gamma_2 D_R^{\gamma_2} & \gamma_3 D_R^{\gamma_3} & \gamma_4 D_R^{\gamma_4} & -\beta_1^B D_R^{\beta_1^B} & -\beta_2^B D_R^{\beta_2^B} & 0 & 0 \\
0 & 0 & 0 & 0 & D_B^{\beta_1^B} & D_B^{\beta_2^B} & 0 & 0 \\
l_1 D_R^{\gamma_1} & l_2 D_R^{\gamma_2} & l_3 D_R^{\gamma_3} & l_4 D_R^{\gamma_4} & 0 & 0 & 0 & 0 \\
l_1 X_B^{\gamma_1} & l_2 X_B^{\gamma_2} & l_3 X_B^{\gamma_3} & l_4 X_B^{\gamma_4} & 0 & 0 & -X_B^{\beta_1^R} & -X_B^{\beta_2^R} \\
l_1 \gamma_1 X_B^{\gamma_1} & l_2 \gamma_2 X_B^{\gamma_2} & l_3 \gamma_3 X_B^{\gamma_3} & l_4 \gamma_4 X_B^{\gamma_4} & 0 & 0 & -\beta_1^R X_B^{\beta_1^R} & -\beta_2^R X_B^{\beta_2^R} \\
X_B^{\gamma_1} & X_B^{\gamma_2} & X_B^{\gamma_3} & X_B^{\gamma_4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & X_R^{\beta_1^R} & X_R^{\beta_2^R}
\end{bmatrix}$$

$$b := \begin{bmatrix}
-A_{B5} + C_3 D_R + C_4 + C_5 D_R^{\gamma_1} + C_6 D_R^{\gamma_2} \\
C_3 D_R + \gamma_1 C_5 D_R^{\gamma_1} + C_6 \gamma_2 D_R^{\gamma_2} \\
-C_3 D_B - C_4 - C_5 D_B^{\gamma_3} - C_6 D_B^{\gamma_4} + \alpha_B \left(D_B y_B + G_B^{unlev}(D_B) \right) \\
-A_{R5} + \alpha_R \left(D_R y_R + G_R^{unlev}(D_R) \right) \\
-A_{R5} + Z\left(X_B\right) \\
X_B Z'\left(X_B\right) \\
-A_{B5} + \hat{d}_B \left(\bar{s} X_B - \frac{K}{y_B} \right) \\
-Z\left(X_R\right) + \hat{d}_R \left(\bar{s} X_R - \frac{K}{y_R} \right)
\end{bmatrix}.$$

Thus, the solution to the remaining unknowns is given by

$$\begin{bmatrix} A_{B1} & A_{B2} & A_{B3} & A_{B4} & C_1 & C_2 & B_1 & B_2 \end{bmatrix}^T = M^{-1} b. \tag{A-99}$$

Case 2: Only 1 regime. Denote the default boundary by D_1, and recall that X_1 is the firm's investment boundary, while $\hat{D}_1$ denotes the default boundary of a firm with only invested assets. Postulating that in the continuation region the required return must be equal to the expected realized return plus the proceeds from debt, we find that the system to solve is:

$$\begin{array}{rcll} d(X) & = & \alpha\left(yX + G^{unlev}(X)\right) & X \leq D_1 \\ rd(X) & = & c + \tilde{\mu}Xd'(X) + \frac{\tilde{\sigma}^2}{2}X^2 d''(X) & D_1 < X < X_1 \\ d(X) & = & \hat{d}\left(\bar{s}X - \frac{K}{y}\right) & X \geq X_1. \end{array} \tag{A-100}$$

The first and second equations are analogous to the two regime case. In the third equation, we postulate that above the exercise boundary X the debt value of the firm must be equal to the one of a firm with only invested assets. As in the two regime case, the conversion of the growth option into assets in place is arranged such that the total value of the firm's assets remains unchanged at the exercise of the option. The boundary conditions are the value-matching conditions at default and exercise:

$$\lim_{X \searrow D_1} d(X) = \alpha\left(yD_1 + G^{unlev}(D_1)\right) \tag{A-101}$$

$$\lim_{X \nearrow X_1} d(X) = \hat{d}\left(\bar{s}X_1 - \frac{K}{y}\right). \tag{A-102}$$

Note that for $X > X_1$ the value of debt is equal to the one of a firm with only invested assets. As the latter is calculated using a no-bubbles condition, we do not have to postulate this condition for the function $d(X)$ again.

The functional form of the solution is

$$d(X) = \begin{cases} \alpha\left(yX + G^{unlev}(X)\right) & X \leq D_1 \\ B_3X^{\beta_1} + B_4X^{\beta_2} + A_5 & D_1 < X < X_1 \\ \hat{d}\left(\bar{s}X - \frac{K}{y}\right) & X \geq X_1, \end{cases} \tag{A-103}$$

where B_3, B_4, A_5, β_1, and β_2 are real-valued parameters to be determined (or to be confirmed). The only region left to solve for is $D_1 < X < X_1$. By plugging the functional form (A-103) into the differential equation (A-100) and comparing coefficients, we find that

$$\beta_{1,2} = \frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2} - \sqrt{\left(\frac{1}{2} - \frac{\tilde{\mu}}{\tilde{\sigma}^2}\right)^2 + \frac{2r}{\tilde{\sigma}^2}} \tag{A-104}$$

$$A_5 = \frac{c}{r}. \tag{A-105}$$

Finally, B_3 and B_4 are determined by the two-dimensional linear system defined by the above boundary conditions:

$$B_3D_1^{\beta_1} + B_4D_1^{\beta_2} + \frac{c}{r} = \alpha\left(yD_1 + G^{unlev}(D_1)\right) \tag{A-106}$$

$$B_3X_1^{\beta_1} + B_4X_1^{\beta_2} + \frac{c}{r} = \hat{d}\left(\bar{s}X_1 - \frac{K}{y}\right). \tag{A-107}$$

Using matrix notation, and

$$
\begin{aligned}
M_1 &:= \begin{bmatrix} D_1^{\beta_1} & D_1^{\beta_2} \\ X_1^{\beta_1} & X_1^{\beta_2} \end{bmatrix}, \\
b_1 &:= \begin{bmatrix} \alpha\left(yD_1 + G^{unlev}\left(D_1\right)\right) - \frac{c}{r} \\ \hat{d}\left(\bar{s}X_1 - \frac{K}{y}\right) - \frac{c}{r} \end{bmatrix},
\end{aligned}
$$

we find that

$$
\begin{aligned}
\begin{bmatrix} B_3 & B_4 \end{bmatrix}^T &= M_1^{-1} b_1 && \text{(A-108)} \\
&= \frac{1}{D_1^{\beta_1} X_1^{\beta_2} - D_1^{\beta_2} X_1^{\beta_1}} \begin{bmatrix} X_1^{\beta_2} & -D_1^{\beta_2} \\ -X_1^{\beta_1} & D_1^{\beta_1} \end{bmatrix} \begin{bmatrix} \alpha\left(yD_1 + G^{unlev}\left(D_1\right)\right) - \frac{c}{r} \\ \hat{d}\left(\bar{s}X_1 - \frac{K}{y}\right) - \frac{c}{r} \end{bmatrix}, && \text{(A-109)}
\end{aligned}
$$

which completes the calculation of the solution.

A.4.2. The valuation of tax benefits

For Case 1, see the main text. Analogously, the value of tax benefits $t(X)$ in Case 2 can be calculated as the value of debt with a recovery rate of zero and a coupon equal to $c\tau$.

A.4.3. The valuation of default costs

Case 1 can be found in the main text. Analogously, the value of default costs $b(X)$ in Case 2 corresponds to the value of debt with a recovery rate of $1 - \alpha$ and a coupon of zero.

A.4.4. Firm value

The main text states the firm value in Case 1. Analogously, for Case 2, the firm value $f(X)$ is

$$f(X) = Xy + G(X) + t(X) - b(X). \tag{A-110}$$

A.4.5. The valuation of equity

Case 1 is given in the main text. For Case 2, the value of equity $e(X)$ is given by

$$e(X) = f(X) - d(X) = Xy + G(X) + t(X) - b(X) - d(X). \tag{A-111}$$

A.4.6. Default policy

Following similar arguments as in Case 1 (main text), the optimal default and investment policies in Case 2, D^* and X^*, are determined by the conditions

$$\begin{cases} e'\left(D^*\right) & = & 0 \\ e'\left(X^*\right) & = & \hat{e}'\left(sX^* - \frac{K}{y}\right) \end{cases} \tag{A-112}$$

A.4.7. Capital structure

Analogously to Case 1 in the main text, denote, for Case 2, by $f^*(X)$ the firm value given ex-post optimal default and expansion thresholds as determined by the system (A-112). The optimal coupon of this firm then solves

$$c^* := \text{argmax}_c f^*(X). \tag{A-113}$$

A.5. The Value of Finite Maturity Debt

We only present the case of the presence of two regimes. The setup and solution for the case of one regime can be derived analogously.

A.5.1. Firms with invested assets only

Hackbarth, Miao, and Morellec (2006) present the solution of a similar model for firms with only invested assets. We consider the standard case that the default boundary in boom is lower than the one in recession, i.e., $\hat{D}_B < \hat{D}_R$. Given debt characteristics (c, m, p), the ODE for the value of debt writes:

For $0 \leq X \leq \hat{D}_B$:

$$\begin{cases} \hat{d}_B(X) & = & \alpha_B X y_B \\ \hat{d}_R(X) & = & \alpha_R X y_R. \end{cases} \tag{A-114}$$

For $\hat{D}_B < X \leq \hat{D}_R$:

$$\begin{cases} \left(r_B^n + m\right)\hat{d}_B(X) & = & c + mp + \tilde{\mu}_B X \hat{d}'_B(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 \hat{d}''_B(X) + \tilde{\lambda}_B\left(\alpha_R X y_R - \hat{d}_B(X)\right) \\ \hat{d}_R(X) & = & \alpha_R X y_R. \end{cases} \tag{A-115}$$

For $X > \hat{D}_R$:

$$\begin{cases} \left(r_B^n + m\right)\hat{d}_B(X) & = & c + mp + \tilde{\mu}_B X \hat{d}'_B(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 \hat{d}''_B(X) + \tilde{\lambda}_B\left(\hat{d}_R(X) - \hat{d}_B(X)\right) \\ \left(r_R^n + m\right)\hat{d}_R(X) & = & c + mp + \tilde{\mu}_R X \hat{d}'_R(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 \hat{d}''_R(X) + \tilde{\lambda}_R\left(\hat{d}_B(X) - \hat{d}_R(X)\right). \end{cases} \tag{A-116}$$

The boundary conditions are the same as in the infinite maturity case, see (A-21)-(A-25). The solution can be found analogously.

Note that for finite maturities, the value of the risk free bond with a given coupon can be calculated as

$$RF = \frac{(c+mp)\left(r_j^n + m + \tilde{\lambda}_i + \tilde{\lambda}_j\right)}{\left(r_i^n + m\right)\left(r_j^n + m\right) + \left(r_j^n + m\right)\tilde{\lambda}_i + \left(r_i^n + m\right)\tilde{\lambda}_j}. \tag{A-117}$$

A.5.2. Firms with invested assets and expansion options

In our framework, debt characteristics (c, m, p) are chosen at initiation and are then constant over time. This setting allows us to calculate the solution for firms with both invested assets and growth options in closed-form, even for finite maturity debt. The standard case with $D_B < D_R$, $\hat{D}_B < \hat{D}_R$, and $X_R > X_B$ is presented. For given debt characteristics (c, m, p), the value of finite maturity corporate debt satisfies the following ODE:

For $0 \leq X \leq D_B$:

$$\begin{cases} d_B(X) &= \alpha_B\left(X y_B + G_B^{unlev}(X)\right) \\ d_R(X) &= \alpha_R\left(X y_R + G_R^{unlev}(X)\right) \end{cases} \tag{A-118}$$

For $D_B < X \leq D_R$:

$$\begin{cases} \left(r_B^n + m\right) d_B(X) &= c + mp + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) \\ & \quad + \tilde{\lambda}_B\left(\alpha_R\left(X y_R + G_R^{unlev}(X)\right) - d_B(X)\right) \\ d_R(X) &= \alpha_R\left(X y_R + G_R^{unlev}(X)\right) \end{cases} \tag{A-119}$$

For $D_R < X < X_B$:

$$\begin{cases} \left(r_B^n + m\right) d_B(X) &= c + mp + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) + \tilde{\lambda}_B\left(d_R(X) - d_B(X)\right) \\ \left(r_R^n + m\right) d_R(X) &= c + mp + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) + \tilde{\lambda}_R\left(d_B(X) - d_R(X)\right) \end{cases} \tag{A-120}$$

For $X_B \leq X < X_R$:

$$\begin{cases} d_B(X) &= \hat{d}_B\left(\bar{s}X - \frac{K}{y_B}\right) \\ \left(r_R^n + m\right) d_R(X) &= c + mp + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) \\ & \quad + \tilde{\lambda}_R\left(\hat{d}_B\left(\bar{s}X - \frac{K}{y_B}\right) - d_R(X)\right) \end{cases} \tag{A-121}$$

For $X \geq X_R$:

$$\begin{cases} d_B(X) &= \hat{d}_B\left(\bar{s}X - \frac{K}{y_B}\right) \\ d_R(X) &= \hat{d}_R\left(\bar{s}X - \frac{K}{y_R}\right). \end{cases} \tag{A-122}$$

Here, $\hat{d}_i(\cdot)$ denotes the value of debt with the same principal, coupon, and debt maturity of a firm with only invested assets, see Appendix A.5.1. The boundary conditions are the same as in the case of infinite maturity debt, see (A-72)-(A-99).

The solution can be derived analogously to the case of infinite maturity debt. Technically, for given debt characteristics (c, m, p), the value of finite maturity debt corresponds to the value of infinite maturity debt with a coupon $c + mp$ and nominal interest rates of $r_i^n + m$.

The assumption that debt is issued at par requires that

$$p = d_i(X), \tag{A-123}$$

where i denotes the regime at initiation. This equation is solved numerically.

A.6. Details on the simulations

A.6.1. Calibration of the idiosyncratic volatility

We calibrate the firm-level idiosyncratic volatility of our BBB sample to the empirically observed total asset volatility of 0.25. The procedure starts by simulating a model-implied economy for 10 years (pre-matching simulation). Next, we match the model-implied distribution after 10 years with the empirical cross-section of BBB-rated firms, and finally simulate the obtained matched sample for another 10 years (post-matching simulation). The average asset volatility of the post-matching simulation is then calculated. The details of this procedure are as follows.

We consider infinite maturity debt in the pre-matching simulation for all debt maturities in the post-matching simulation. We do so to abstract away from the impact of different initial principals on the results, allowing us to analyze the pure effect of debt maturities on credit spreads in the post-matching simulation. Additionally, starting with infinite maturity debt yields initial leverage ratios (principals) close to the ones empirically reported.[32] The model-implied economy is generated as follows. Starting with a value firm ($s = 0$), we generate a range of firms by increasing the option scale parameter s by steps of 0.05, up to the largest possible value of s such that the option is not exercised immediately. At initiation, the capital structure is chosen optimally for all firms. For each option scale parameter s, 50 firms are considered, resulting in an initial sample of more than 3,000 firms. During the 10-year pre-matching simulation, firms default and expand optimally. Defaulted firms are not replaced, and exercised firms continue as firms with only invested assets. At the end of the pre-matching simulation, we calculate the model-implied leverage and asset composition ratio for each firm, using the assumed debt maturity and the corresponding optimal boundaries. We obtain a model-implied distribution of firms covering a broad range of both asset composition ratios and leverage ratios.

In the second step, we match our average historical distribution of BBB-rated firms with its model-implied counterpart. For each observation in the average historical distribution, we select the firm in our model-implied economy at the final period of the pre-matching simulation which exhibits the minimum distance regarding the percentage deviation from the target market leverage and asset composition ratio. That is, the empirical observation of a firm with leverage lev_{emp} and asset composition ratio acr_{emp} is matched with the model-implied firm with leverage lev_{mi} and asset composition ratio acr_{mi} if - given the set of all model-implied firms - it minimizes the Euclidean distance

$$\sqrt{\left(\frac{lev_{emp} - lev_{mi}}{lev_{emp}}\right)^2 + \left(\frac{acr_{emp} - acr_{mi}}{acr_{emp}}\right)^2}. \tag{A-124}$$

[32] A robustness analysis confirms that starting with finite maturity debt in the pre-matching simulation yields slightly lower credit spreads in the post-matching simulation, as the initial principals are smaller.

The final step conducts a post-matching simulation with the obtained sample of model implied BBB-firms over 10 years. For each simulation, we obtain the realized asset volatility for each firm, and calculate the resulting average asset volatility over firms. When measuring and averaging asset volatilities, we incorporate the entire initially matched BBB-sample, including the evolution of the assets of firms which default during the 10-year post-matching simulation. This approach avoids a weighting bias when averaging over simulations towards firms with lower leverage and asset volatility which have a smaller tendency to default during the post-matching simulation.

The pre-matching simulation and the subsequent matching is conducted 20 times. The initial regime is chosen according to the stationary distribution of the states, i.e., the pre-matching simulation starts in boom $100\frac{\lambda_R}{\lambda_B+\lambda_R}\%$ of the total number of simulations. This approach also guarantees convergence to the steady-state distribution of regimes at the time of matching. For each matched sample of firms, the post-matching simulation is run 50 times. These numbers result in a total of 1,000 simulations. The procedure is conducted for different post-matching debt maturities.

A.6.2. Simulation of the true cross-section

To ensure consistency, the simulation of the true cross-section is done analogously to the one performed to calibrate the idiosyncratic volatility: We first simulate a model-implied distribution of firms for 10 years (pre-matching simulation), and then match the model-implied distribution with the average empirical cross-section (for details, see above). The final step consists of simulating the matched sample for 20 years (post-matching simulation). We assume that firms default and exercise optimally. Defaulted firms are immediately deleted, whereas exercised firms are maintained in the sample, and continue as firms with only invested assets. Credit spreads and leverage ratios are measured during 5 years after the matching: For each firm in the sample, we calculate the actual credit spread and leverage every month, and then report the average over all firms and all simulations. Default rates are observed for 5, 10, and 20 years. In order to incorporate the impact of the realized regimes at initiation and at the time of matching, we present quantiles of post-matching average rates. As in the calibration of the volatility, the initial state is chosen according to the stationary distribution. The pre-matching simulation is run 20 times, and the post-matching simulation is conducted 50 times, resulting in a total of 1,000 simulations.

A.7. Robustness tests

A.7.1. Financing the exercise of the growth option by issuing additional equity

We consider the case that the exercise price λ of the growth option is financed by issuing additional equity. The corresponding system of ODEs for corporate debt is:

For $0 \leq X \leq D_B$:

$$\left\{\begin{array}{lcl} d_B(X) & = & \alpha_B\left(Xy_B + G_B^{unlev}(X)\right) \\ d_R(X) & = & \alpha_R\left(Xy_R + G_R^{unlev}(X)\right) \end{array}\right. \tag{A-125}$$

For $D_B < X \leq D_R$:

$$\begin{cases} r_B^n d_B(X) &= c + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) + \tilde{\lambda}_B \left(\alpha_R \left(X y_R + G_R^{unlev}(X)\right) - d_B(X)\right) \\ d_R(X) &= \alpha_R \left(X y_R + G_R^{unlev}(X)\right) \end{cases} \tag{A-126}$$

For $D_R < X < X_B$:

$$\begin{cases} r_B^n d_B(X) &= c + \tilde{\mu}_B X d_B'(X) + \frac{1}{2}\tilde{\sigma}_B^2 X^2 d_B''(X) + \tilde{\lambda}_B (d_R(X) - d_B(X)) \\ r_R^n d_R(X) &= c + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) + \tilde{\lambda}_R (d_B(X) - d_R(X)) \end{cases} \tag{A-127}$$

For $X_B \leq X < X_R$:

$$\begin{cases} d_B(X) &= \hat{d}_B(\bar{s}X) \\ r_R^n d_R(X) &= c + \tilde{\mu}_R X d_R'(X) + \frac{1}{2}\tilde{\sigma}_R^2 X^2 d_R''(X) + \tilde{\lambda}_R \left(\hat{d}_B(\bar{s}X) - d_R(X)\right) \end{cases} \tag{A-128}$$

For $X \geq X_R$:

$$\begin{cases} d_B(X) &= \hat{d}_B(\bar{s}X) \\ d_R(X) &= \hat{d}_R(\bar{s}X) \,. \end{cases} \tag{A-129}$$

The boundary conditions now read:

$$\begin{aligned} \lim_{X \searrow D_R} d_B(X) &= \lim_{X \nearrow D_R} d_B(X) \\ \lim_{X \searrow D_R} d_B'(X) &= \lim_{X \nearrow D_R} d_B'(X) \\ \lim_{X \searrow D_B} d_B(X) &= \alpha_B \left(D_B y_B + G_B^{unlev}(D_B)\right) \\ \lim_{X \searrow D_R} d_R(X) &= \alpha_R \left(D_R y_R + G_R^{unlev}(D_R)\right) \\ \lim_{X \searrow X_B} d_R(X) &= \lim_{X \nearrow X_B} d_R(X) \\ \lim_{X \searrow X_B} d_R'(X) &= \lim_{X \nearrow X_B} d_R'(X) \\ \lim_{X \nearrow X_B} d_B(X) &= \hat{d}_B(\bar{s}X_B) \\ \lim_{X \nearrow X_R} d_R(X_R) &= \hat{d}_R(\bar{s}X_R) \,. \end{aligned} \tag{A-130}$$

The solution to this system follows by standard arguments from the theory of differential equations. Technically, this modification constitutes a simplification of the presented main case: The functional form is straightforward and does not need to be determined as the solution of an inhomogeneous ODE using the fundamental system of solutions of the homogenous ODE (cf. footnote 31). Therefore, we do not present the solution here.

The Impact of Managerial Control over Cash on Credit Risk and Financial Policy

Marc Arnold*

October 14, 2011

Abstract

This article analyzes the impact of managerial control over cash holdings on credit risk and corporate financial policy. I identify two channels which drive the total effect of cash on credit risk. First, the fact that a manager can use cash to service debt when equityholders are unwilling to inject funds into the firm reduces credit risk. Second, as equityholders anticipate that ceasing to inject funds does not lead to immediate default, they optimally stop contributing funds earlier in firms with larger cash holdings. The second channel increases credit risk. Because a manager holds excess cash to reduce the probability of corporate default, the relative strength of the two opposing channels can explain the relationship between observed excess cash levels and firm or industry characteristics. The paper also discusses the impact of the market for corporate control, the agency costs of cash, and the under-leverage puzzle.

*Arnold is a Ph.D. student at the Swiss Finance Institute at the University of Zurich. Office address: Institute of Banking and Finance, Plattenstrasse 14, CH-8032 Zurich, Switzerland, Phone: 0041-44-634-3957, Email: marc.arnold@bf.uzh.ch. I thank Elisabeth Megally, Jules Munier, Kjell Nyborg, Jean-Charles Rochet, Tatjana-Xenia Puhan, Cornelius Schmidt, René Stulz, Alexander Wagner, Ramona Westermann, and seminar participants at the Research Day of the NCCR FINRISK for helpful comments. This research was supported by the NCCR FINRISK, the Swiss Finance Institute, and the ProDoc, a research instrument of the Swiss National Science Foundation.

1. Introduction

The empirical corporate finance literature finds that firms accumulate excess cash if they have the opportunity to do so. According to Dittmar, Mahrt-Smith, and Servaes (2003), and Pinkowitz, Stulz, and Williamson (2003), for example, managers are more likely to build excess cash balances when they are entrenched, or when shareholder rights are low. There also is anecdotal evidence that managers target cash in excess of the level desired by equityholders. When Kirk Kerkorian attempted a takeover of Chrysler Corporation, Chrysler's management insisted on keeping large cash reserves because it considered drawing funds from alternative sources during a financial crisis unrealistic (Yun 2009). Similarly, Froot (1992) describes the case where Intel's executives defended their massive cash balance by saying that it was an important competitive weapon, especially during recession. Given the tremendous amount of cash in firms' balance sheets reported in Bates, Kahle, and Stulz (2009), and the observation in Faulkender and Petersen (2006), and Pinkowitz, Stulz, and Williamson (2006) that the marginal value of cash to equityholders is considerably below one, it seems important to understand the sources and consequences of excess cash.

Managers state that, for them, cash most importantly serves a self-preservation motive, i.e., a basic financial insurance function to provide a buffer against corporate bankruptcy in future hard times (Lins, Servaes, and Tufano 2009). The present paper shows that this motive can explain observed excess cash levels. It also allows to link excess cash to firm and industry characteristics. Additionally, I analyze the impact of excess cash on credit risk and on the corporate financial policy.

I extend the basic trade-off model of capital structure in the spirit of Leland (1994) by incorporating both a corporate cash policy and agency conflicts between managers and shareholders. Shareholders agree to leave cash within the firm because of future profitable investment opportunities which can only be captured by investing quickly out of cash. This is the standard precautionary motive (Huberman 1984, Kahan and Yermack 1998, Mikkelson and Partch 2003). Due to their control over cash resources, however, managers have an additional motive for holding cash: Cash reduces a firm's default risk by providing a buffer against bankruptcy during times where equity financing is not available. It is shown that due to this self-preservation motive, managers target a cash level in excess of the one which maximizes equityholders' wealth when they trade off the impact of cash on the value of their fixed income stream against the impact on their equity compensation. They can target excess cash because of costs to replace the management.

To explain to what extent excess cash is driven by managers' self-preservation motive, one first needs to understand how cash affects the default risk of firms. I identify two channels. First, as managers control cash, they can use it to defer default when equityholders are unwilling to inject funds during economic distress. Postponement of default is possible by servicing debt out of cash holdings until cash is exhausted. Using cash to defer default is beneficial for managers because it allows them to obtain the fixed salary, and their personal utility from investing cash over an extended period of time. This result reflects the common intuition about how cash reduces credit risk.

Second, I show that there also is an indirect effect of cash on credit risk. In standard capital structure models without cash, it is optimal for equityholders to inject funds until the asset value reaches a certain trigger. As equityholders are protected on the downside due to limited liability, but still participate on the upside if the asset value recovers, the trigger is below the nominal value of debt. Hence, the option feature of equity induces access to equity funds even if the firm is insolvent on a flow basis. In a firm with cash holdings, equityholders anticipate that ceasing to inject funds will not lead to immediate default. In case of insolvency on a flow basis, they can stop injecting funds without losing their option claim on the firm's assets because debt will still be serviced out of cash holdings. The possibility to save coupon payments before a potential default induces equityholders to stop contributing funds earlier than in a firm without cash. Hence, cash makes it harder for firms to obtain equity funds during bad times. This important indirect effect of cash increases credit risk. It has, so far, been neglected in the theoretical literature.

I then quantitatively investigate the relative strength of the direct and indirect effects of cash holdings on credit risk. I find that excess cash due to managers' self-preservation motive can be substantial. In the baseline firm, managers' target cash level is about 20% above the one which maximizes shareholder wealth. Depending on firm and industry characteristics, excess cash can reach 182% of the level that shareholders would deem optimal. The agency costs of this distortion are important, implying up to a 9.7% wealth loss to equityholders for realistic parameters. The bias towards excess cash is restricted by the equityholders' willingness to contribute the idle cash, and by the market for corporate control.

I also employ a dynamic approach in the spirit of Bhamra, Kuehn, and Strebulaev (2010) by simulating over time a model-implied cross-section of firms which is structurally similar to S&P 500 firms. Importantly, it is shown that the self-preservation motive induces an amount of excess cash which generates realistic cash-properties in an economy. In particular, I find that the average market value of cash to firms, and the marginal value of cash to equityholders in a simulated economy correspond closely to their empirical counterparts reported in Pinkowitz and Williamson (2002) and Faulkender and Petersen (2006), respectively.

The presented model derives new predictions, and also talks to recent empirical findings. First, I show that when incorporating managerial control over cash, managers use cash to service debt instead of investment during economic distress. Hence, close to default, cash does not generate value to equityholders but mainly accrues to debtholders. Equityholders should, therefore, optimally leave less cash within riskier firms. Empirically, however, riskier firms hold larger amounts of cash (Acharya, Bharath, and Srinivasan 2007, Bates, Kahle, and Stulz 2009). I show that the self-preservation motive can explain this relationship: Managers of risky firms optimally target larger excess cash than those of safe firms because the direct effect of cash on credit risk is positively related to the asset volatility.

Second, a number of recent papers point to a hedging motive behind corporate cash holdings, as cash may allow financially constrained firms to hedge future investment against income shortfalls (Acharya, Almeida, and Campello 2007, Faulkender and Petersen 2006, Acharya, Bharath, and Srinivasan 2007, Denis and Sibilkov 2010). I argue that when agency conflicts are incorporated, cash is not a suitable hedging tool to ensure investment in future bad states: Managers will use these funds to service debt payments instead of investing when equityholders stop contributing funds.

Finally, managerial control over cash also affects the choice of optimal leverage. Besides providing a tax shield and imposing bankruptcy costs, debt also causes underinvestment costs because cash is used to service debt instead of being invested during economic distress. These costs induce lower optimal leverage than in the standard trade-off theory of capital structure.

1.1. Related literature

Understanding what triggers default is a central issue in the literature on capital structure and credit risk. Models of economic distress assume that firms default when the market value of assets falls below a certain boundary (Leland 1994, Longstaff and Schwartz 1995). Should a temporary cash shortfall lead to a liquidity crisis, shareholders will meet the required debt payments by raising outside financing. Cash holdings are, therefore, irrelevant under this approach. In contrast, financial distress models postulate that default occurs when cash shortages result in the inability of firms to make promised debt payments (Asquith, Gertner, and Scharfstein 1994). Due to market frictions, a distressed firm may be unable to raise necessary external financing, despite the fundamentally sound nature of its business.

It is surprising that even though holdings of cash are extensively used in empirical bankruptcy-predicting models (Altman 1968, Zmijewski 1984), no theoretical paper exists that rigorously investigates the interaction between economic and financial distress. My work combines the two approaches by incorporating cash in the standard economic distress model of Leland (1994). In particular, it is shown that if firms have both access to outside financing and cash, they default when the asset value remains below the threshold where equityholders optimally stop contributing funds, until cash is exhausted. This setting reflects the empirical finding in Davydenko (2007) that a transitory cash shortage can trigger default only to the extent that a firm is restricted from raising new financing against its remaining assets. The authors show that persistent economic distress causes financial distress, when money-losing firms run out of the liquid funds necessary to pay their creditors and suppliers. While this basic point is already recognized in the empirical literature, the implications from a realistic default triggering of firms with cash for credit risk and financial policy have not been fully developed. My model also aims at filling this gap.

The presented model is closely related to the growing literature on the role of cash holdings within structural models. Gamba and Triantis (2008) explain how debt flotation costs lead to simultaneous cash and debt holdings. Morellec and Nikolov (2009) find that the secular increase in cash holdings in the US can partially be explained by the intensity of industry competition, and Hugonnier, Malamud, and Morellec (2010) show how capital market supply frictions affect corporate behavior. While these papers abstract away from both agency issues and credit risk, there exist two closely related studies on cash holdings which also incorporate agency. Nikolov and Whited (2009) estimate a dynamic model of firm investment and cash accumulation in the presence of shareholder-manager conflicts. The authors argue that when management derives a psychological flow benefit per unit of installed capital, managers like to build empires, hold suboptimal levels of internal liquidity, and use too much costly external financing. While I also consider private utility from investment, my work incorporates credit risk which induces managers to target excess cash to protect their fixed compensation and future investment against bankruptcy. In the model of Mahmudi and Pavlin (2010), the objective function of the manager is extended to capture a perceived cost to payout reductions. Higher cash today decreases the probability of having to cut dividends tomorrow, and, consequently, lowers managers' expected disutility. My model is similar to theirs with respect to the idea that managers' personal motive for saving cash can be important for firm policy. Instead of perceived costs of payout reductions, however, the reason for excess cash in my model is the self-preservation motive induced by credit risk. In short, the main difference of this paper to the existing theoretical work is that I simultaneously capture corporate cash, risky debt, and agency issues in an integrated framework.

My work also contributes to the discussion in the empirical literature on the relation between cash holdings and agency problems by showing that self-preservation is an important motive behind excess cash. It is widely accepted that firms hold dramatically more cash than one would expect from the static tradeoff theory where managers maximize shareholders' wealth (Opler, Pinkowitz, Stulz, and Williamson 1999). Additionally, Dittmar, Mahrt-Smith, and Servaes (2003), Pinkowitz, Stulz, and Williamson (2006), Kalcheva and Lins (2007), and Harford, Mansi, and Maxwell (2008) show that the market value of cash reserves is lower when firms are poorly governed, and when there is weak shareholder protection. While they argue that cash hoarding by managers is value-reducing, Mikkelson and Partch (2003) find that a policy of high cash holdings may in fact be value-enhancing. All these studies, however, focus primarily on the effects rather than on the motive behind excess cash. An exception is the work of Bates, Kahle, and Stulz (2009) who show that riskier cash flows, particularly concentrated among smaller high tech firms, have caused the dramatic growth of excess cash in recent years.

The remainder of the paper is organized as follows: Section 2 introduces the model. Section 3 and 4 discuss the results and empirical implications. Model extensions are presented in Section 5 before I conclude in Section 6.

2. The model

The model setup is a standard continuous time model of capital structure decisions in the spirit of Leland (1994), extended to account for cash holdings. All agents including managers are risk-neutral and discount at a constant interest rate r. Assets are continuously traded in arbitrage-free markets. Time is continuous, and uncertainty is modeled by a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{\{t\geq 0\}}, \mathbb{Q})$. The firm issues perpetual debt with contractual coupon payments c. The motive for issuing debt is the tax shield provided by debt financing. Corporate taxes are paid at a rate τ, and full offsets of corporate losses are allowed. The value of the firm's invested assets $(V_t)_{t\geq 0}$ is assumed to be independent of capital structure choices and the cash policy. Under the risk-neutral measure $\mathbb{Q}$, it is governed by the following process:

$$dV_t = (r - \delta)V_t dt + \sigma V_t dW_t \tag{1}$$

In Equation (1), δ is the firm's payout ratio from invested assets, σ denotes the constant volatility of asset returns, and $(W_t)_t \geq 0$ is a standard Brownian motion. The model abstracts away from any transaction costs such as equity issuance costs.

Besides invested assets, a firm also holds cash L if it is worthwhile for equityholders to do so. As in Holmstroem and Tirole (2000), cash is defined as holdings which the firm can quickly resell or pledge as collateral at its face value. Managers raise additional equity in order to obtain the cash when initiating the firm. Following Keynes (1936), numerous empirical studies suggest that the precautionary motive is a critical determinant of corporate cash holdings (Harford, Mikkelson, and Partch 2003, Mikkelson and Partch 2003, Almeida, Campello, and Weisbach 2004, Faulkender and Petersen 2006, Haushalter, Klasa, and Maxwell 2007).[1] To capture this motive, I assume that firms expect to have access to profitable investment opportunities in the future which can only be captured with cash that is readily available, and are not feasible with outside financing. The literature describes several reasons for this friction.[2] The investment opportunities occur periodically according to a Poisson process with intensity y. Each investment produces a profit equal to $R(L) = AL^{\gamma}$, with $A > 0$, and $0 < \gamma < 1$. For tractability, it is assumed that these profits accrue instantaneously, and that managers invest the entire amount L.[3] The profit function R has

[1] The basic idea and the derived insights on excess cash also apply to alternative motives of corporations to hold cash. The only necessary assumption is that cash holdings cause benefits which occur continuously or periodically, such as reductions in competition or transaction costs (Miller and Orr 1966, Baskin 1987, Froot 1992, Morellec, Nikolov, and Schürhoff 2009), or benefits described in the real options literature. The latter argues that cash allows managers to flexibly and rapidly adapt to altered future conditions which can expand a project's or investment opportunity's value (Trigeorgis 1993, Yeo and Fasheng 2003).

[2] Borrowing or issuing stock entails direct and indirect costs such as the effects of conflicts between bondholders and stockholders described in Myers (1977) or Jensen (1986), and the effects of information problems with outside investors identified by Myers and Majluf (1984). With cash, firms avoid forgoing profitable investment opportunities when it becomes too costly to raise external money for certain projects (Huberman 1984, Lins, Servaes, and Tufano 2009). Similarly, an attempt to convince the market of a project can take time and entails the dissemination of information, which may deteriorate the project's value, because public information can be used by competitors (Bhattacharya and Ritter 1983).

[3] In case of a time lag between investment and the profit, one needs to specify what happens when economic distress occurs during the lag. If managers can pledge or sell the investment, the predictions of my model remain unchanged. Assuming investment of the entire amount is natural in the current setup: Only the precautionary motive induces equityholders to leave any idle cash within the firm. As $R(L)$ is constant, managers have no justification for keeping unused cash within the firm.

standard properties, i.e., it is increasing, concave, and continuously differentiable.[4] The net payout to equityholders from the entire cash L being invested upon the occurrence of an opportunity is equal to $R(L) - L$. L is deducted because the initial cash level is maintained. Hence, the total payout of L units of cash periodically invested up to time t can be expressed as a compound Poisson process, i.e., as

$$I_t = \sum_{i=1}^{N_t} (R_{t_i}(L) - L), \tag{2}$$

where N_t is a pure Poisson process with intensity y, and $R_{t_i}(L)$ represent the periodical profits from the cash amount L invested at time t_i. The intensity is first modeled to be independent of V. Later, I consider the case where it is stochastic, and correlated with V.

While the ability to capture investment opportunities is beneficial for firms, holding cash is also costly. In particular, the continuous yield l from cash holdings is assumed to be smaller than the interest rate r, expressing the notion that cash remains invested in very liquid assets which typically yield a low return. The relation $l < r$ may also capture the tax disadvantage of cash holdings (Graham 2000, Riddick and Whited 2009).

Even though a firm may operate forever, shareholders have limited liability and, thus, the option to default on their obligations. In the framework of Leland (1994) and Duffie and Lando (2001), bankruptcy is triggered when they cease injecting funds into the firm. Shareholders will optimally do so when the economic net worth of equity is zero. This stock-based definition of default implies that the firm is insolvent on a flow basis at the default date. Managers can not operate the firm's assets beyond the default date which maximizes equity value because they are unable to meet current debt obligations after this date. Hence, shareholders have control rights over the decision to default. This existing framework neglects the impact of cash holdings on the default policy. Cash introduces an important source of manager flexibility: Managers' control over cash holdings, or "deep pockets", allows them to service debt out of cash holdings when equityholders cease to inject the necessary capital to finance continued operation of the firm. In particular, I assume that they can defer default by pledging the cash (L) to debtholders in order to satisfy the coupon payments c up to L/c units of time. Managers will always pledge when equityholders cease to inject funds because they lose any claim on the firm if it defaults.

To formalize this default setting, I define the following random variables:

[4] The most natural interpretation of the assumed concavity is that there are technologically decreasing returns to scale, or a downward sloping demand curve. The qualitative results do not critically depend on the supposed functional form of the profit function.

$$g_t^{V_B} = sup\big(s \leq t : V_s = V_B\big), \tag{3}$$

$$\theta^{V_B} = inf\big(t \geq 0 : (t - g_t^{V_B}) \geq d, V_t \leq V_B\big), \tag{4}$$

where V_B is the threshold value of invested assets below which equityholders cease injecting funds into the firm, $g_t^{V_B}$ denotes the last time before t that the value of the firm's assets reaches V_B, $d = \frac{L}{c}$, and θ^{V_B} denotes the liquidation time. As long as $V < V_B$, the firm is in an "economic distress" situation because it is unable to obtain outside equity financing. I assume that distress-costs (ρV) accrue during economic distress, and that $\rho < \delta$.[5] Distress costs reflect the idea that firms incur indirect costs such as lost profits when customers switch to other producers in anticipation of a potential liquidation (Webb 1987). As dividends are given by $\delta V - c$, and $(\delta - \rho)V$ outside and inside economic distress, respectively, this setting is in accordance with the observation that dividends are often cut but not omitted during distress (DeAngelo and DeAngelo 1990).

The default time θ^{V_B} is finally given by the first time the pledged cash amount L is exhausted, i.e., does not cover additional coupon payments. It occurs when the value of the firm's assets has spent $d = \frac{L}{c}$ units of time consecutively below the default threshold V_B. This default definition captures in a simple way the notion in Davydenko (2007), who finds empirical evidence that persistent economic distress leads to default when firms run out of liquid funds to meet debt service payments. Default induces immediate liquidation where debtholders recover the value of the unlevered assets minus the direct bankruptcy costs. The latter correspond to a fraction $(1 - \alpha)$ of the unlevered asset value. Proceeds from liquidation are distributed according to absolute priority. Additionally, debtholders receive the pledged cash amount which represents the (delayed) debt service during economic distress.[6]

Note that cash has two functions in my model: It is either used for investments, or pledged to debtholders to avoid immediate default during economic distress. This setting is in accordance with evidence from the empirical literature. Campello, Graham, and Campbell (2010) show that the recent financial crisis led firms to burn through their cash to finance continued operation, and

[5] If $\rho > \delta$, d becomes stochastic because not only c but also $(\rho - \delta)V$ has to be financed out of cash holdings during economic distress.

[6] Alternative debt service settings during times where equityholders cease to inject funds can affect the valuation of debt. For example, direct coupon payments out of cash holdings slightly increase the value of debt because cash already accrues before default via these payments, whereas with pledging, debtholders recover the entire cash at default. The quantitative effect is, however, limited and has no impact on the basic prediction of the model.

to bypass attractive investment opportunities. Opler, Pinkowitz, Stulz, and Williamson (1999) find that firms typically lose excess cash by covering losses in weaker financial conditions, rather than by spending on new projects. Additionally, common bond covenants described by Smith and Warner (1979) induce a similar setting even if cash is not physically pledged to debtholders in reality because they restrict investment or spending cash if key financial ratios are below a certain threshold.

I assume that managers face a standard compensation package with fixed income ω, and a variable compensation. The underlying idea of incorporating a fixed income is that managers earn less after losing their job (Fee and Hadlock 2004). As in Nikolov and Whited (2009), managers' variable compensation consists of a share Ψ of the equity capital.[7] Following the spending hypothesis of Jensen (1986), I also incorporate the notion that managers derive utility from investing cash reserves, i.e., from undertaking new projects. Their private utility obtained from investing the amount L is assumed to correspond to $\eta log(L)$.[8] Taking the logarithm reflects the idea that managers' marginal utility from investments is positive and decreasing in the spent amount. It captures in a simple fashion the notion reported in the literature that firms invest more when their cash holdings are high (Almeida, Campello, and Weisbach 2004, Kim, Mauer, and Sherman 1998, Marchica and Mura 2008). For simplicity, the benefits a manager receives from ω, Ψ, and η are assumed to be additively separable. Additionally, as the model is set up in a risk-neutral framework, I abstract away from deducting the value of management's compensation package from the asset value.

In a standard neoclassical model, firms target a cash level which maximizes the equity value by balancing marginal benefits against marginal costs of cash. Inspired by the empirical literature around Opler, Pinkowitz, Stulz, and Williamson (1999), Faulkender and Petersen (2006), and Dittmar and Mahrt-Smith (2007), I incorporate that entrenched managers deviate from this level, and target a corporate cash level which maximizes their own private benefits. They can do so because replacing the management by a takeover is costly. Takeover threats are modeled by defining an alternative management team A which would select the optimal cash policy of equityholders, and takeover costs T. The intensity of the occurrence of investment opportunities under the alternative management team (ability to find profitable investment opportunities) is y^A.

[7] The survey of Murphy (1999) on CEO compensation documents that compensation packages typically consist of a fixed wage, profit share, straight equity, and options. I ignore the profit share because cash has limited effect on firm profits. The model could easily be adapted to separate between straight equity and options. I do not expect new insights from analyzing option based compensation because cash only marginally changes firm volatility.

[8] Alternatively, one could introduce a psychological flow benefit for managers per unit of capital, or per unit of cash under their control, expressing managerial preference for empire building. Untabulated results show that the qualitative predictions maintain in such a setting.

Recent research aims at explicitly tracking the accumulation, or build-up, process of cash holdings (Décamps and Villeneuve 2007, Bolton, Chen, and Wang 2009, Riddick and Whited 2009). In contrast to this literature, I argue that firms - much like a debt policy - initiate a cash policy characterized by their target cash level when setting up the firm. In fact, Opler, Pinkowitz, Stulz, and Williamson (1999) and Ozkan and Ozkan (2004) indicate that firms target a particular cash level, and (Bates, Kahle, and Stulz 2009) show that they seem to initiate cash holdings at their IPO instead of building-up cash over time. For tractability, I assume that firms maintain this target cash level as long as they have access to financing, i.e., are outside economic distress. In reality, of course, firms' cash holdings are not constant. If they only deviate randomly from their target cash levels, however, one should be able to explain empirically observed patterns by studying the cash policy.

3. Results

3.1. Valuation of corporate securities

Before analyzing the impact of cash holdings on credit risk and financial decisions, it is useful to identify the sources of value within the firm. The value of a leveraged firm with cash holdings can be written as:

$$\begin{aligned} F(V,L) = \mathbb{E}_Q\Big(& \int_0^{\theta^{V_B}} e^{-ru}[\delta V_u + \tau c 1_{V_u > V_B} - \rho V_u 1_{V_u < V_B}]du\Big) \\ & + \mathbb{E}_Q\Big(\sum_{i=1}^{N_{\theta^{V_B}}} e^{-rt_i}(R_{t_i}(L) - L) 1_{V_u > V_B}\Big) + lL\mathbb{E}_Q\Big(\int_0^{\theta^{V_B}} e^{-ru} 1_{V_u > V_B} du\Big) \\ & \quad + \alpha \mathbb{E}_Q(e^{-r\theta^{V_B}} V_{\theta^{V_B}}) + L\mathbb{E}_Q(e^{-r\theta^{V_B}}). \end{aligned} \tag{5}$$

The first term on the right hand side captures the cash flows of assets in place until liquidation, i.e., the payout (δ), tax benefits akin to a security that pays a constant coupon (τc), and economic distress costs (ρ). The second term is the value of the expected payouts from investing cash in the periodically occurring opportunities, and the third term is the value of the continuous yield from cash holdings. Note that these benefits of cash only accrue as long as the firm is not in economic distress. Below V_B, management pledges the cash to secure coupon payments in order to defer default. Strictly speaking, the loss during economic distress of both the opportunity to invest and the yield from cash also reflects (indirect) distress costs. The fourth term on the right hand side of Equation (5) captures the recovery value upon liquidation. The last term, finally, expresses the notion that the (pledged) cash accrues to debtholders at default. Solving this equation yields the following proposition.

Proposition 1. *The firm value for a given default threshold satisfies*

$$\begin{aligned}
F(V,L) = V &- \frac{\delta V_B}{\lambda^2+(\sigma+b)\lambda}\Big(\frac{V}{V_B}\Big)^{-\xi} - \delta\frac{V_B}{\lambda}\Big(\frac{V}{V_B}\Big)^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma-b} \\
&+\frac{\tau c+\Gamma(L)}{r}\Big(1-\Big(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\Big)\Big(\frac{V}{V_B}\Big)^{-\xi}\Big) \\
&+\alpha V_B\Big(\frac{V}{V_B}\Big)^{-\xi}\frac{\Phi(-(b+\sigma)\sqrt{d})}{\Phi(\lambda\sqrt{d})}+L\frac{e^{a\sqrt{2r}}}{\Phi(\sqrt{2rd})} \\
&+(\delta-\rho)\Big((1/\delta)\Big(V-\frac{\phi(-(\sigma+b)\sqrt{d})}{\phi(\lambda\sqrt{d})}V_B\Big(\frac{V}{V_B}\Big)^{-\xi}\Big) \\
&-\Big(V/\delta-\frac{V_B}{\lambda^2+(\sigma+b)\lambda}\Big(\frac{V}{V_B}\Big)^{-\xi}-\frac{V_B}{\lambda}\Big(\frac{V}{V_B}\Big)^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma-b}\Big)\Big),
\end{aligned} \tag{6}$$

where $\Phi(x) = 1 + x\sqrt{2\pi}exp(\frac{x^2}{2})N(x)$, *N is the standard normal cumulative distribution function,* $\lambda = \sqrt{2r+b^2}$, $a = (1/\sigma)(log(V_B/V))$, $\xi = (1/\sigma)(b+\lambda)$, *b is defined as* $b = (1/\sigma)(r-\delta-\sigma^2/2)$, *and* $\Gamma(L) = y(R(L)-L)+lL$.

Proof. See in the Appendix. □

Next, the value of debt is derived. I consider debt contracts which are characterized by a perpetual flow of coupon payments (c), and the commitment that the firm is immediately liquidated if it is unable to service debt. As equityholders stop injecting funds into the firm when V reaches V_B, management needs to pledge the available cash to debtholders in order to continue satisfying the debt service during economic distress.

Prior to calculating the value of corporate debt, I need to define what happens if the asset value recovers from economic distress before cash is exhausted, i.e., if V reaches V_B between $g_t^{V_B}$ and θ^{V_B}. Suppose that renegotiation takes place in this case whereat debtholders try to induce equityholders to resume the debt service, and that a failure to renegotiate triggers immediate default. By construction of the economic distress threshold, equityholders are indifferent about resuming the debt service if the initial cash level (L) is maintained when V reaches V_B after $g_t^{V_B}$. Hence, debtholders will always waive the pledged coupon payments for successful renegotiation, i.e., leave L units of cash within the firm, if[9]

$$D(V_B, L) \geq \alpha V_B + L. \tag{7}$$

The left hand side of Inequality (7) denotes the continuation value of debt at the economic distress threshold (V_B). The right hand side is the recovery value if renegotiation fails, and default is triggered. Given that Inequality (7) holds, the value of corporate debt D(V,L) can be defined as

$$D(V, L) = E_Q(\int_0^{\theta^{V_B}} e^{-ru} c 1_{V_u > V_B} du) + \alpha E_Q(e^{-r\theta^{V_B}} V_{\theta^{V_B}}) + L\mathbb{E}_Q(e^{-r\theta^{V_B}}). \tag{8}$$

The first term on the right hand side captures the regular coupon payments up to economic distress. The second term is the recovery value of the firm. Note that debtholders anticipate the occurrence of economic distress, default, and the outcome of renegotiation. The last term in Equation (8) expresses the value of the pledged cash to debtholders. Proposition 2 shows the debt value and the renegotiation condition in closed form.

Proposition 2. *The debt value of a firm is given by*

$$\begin{aligned} D(V, L) = {} & \frac{c}{r}(1 - (\frac{\lambda - b}{2\lambda} + \frac{\lambda + b}{2\lambda} \frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{V}{V_B})^{-\xi}) \\ & + \alpha V_B (\frac{V}{V_B})^{-\xi} \frac{\Phi(-(b+\sigma)\sqrt{d})}{\Phi(\lambda\sqrt{d})} + L \frac{e^{a\sqrt{2r}}}{\Phi(\sqrt{2rd})}. \end{aligned} \tag{9}$$

Debt is successfully renegotiated at V_B after $g_t^{V_B}$, if

$$\begin{aligned} & \frac{c}{r}(1 - \frac{\lambda - b}{2\lambda} - \frac{\lambda + b}{2\lambda} \frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}) \\ & + \alpha V_B (\frac{\Phi(-(b+\sigma)\sqrt{d})}{\Phi(\lambda\sqrt{d})} - 1) + L(\frac{1}{\Phi(\sqrt{2rd})} - 1) \geq 0. \end{aligned} \tag{10}$$

Proof. See in the Appendix. □

[9] The sharing rule for the cash flows from assets in place at renegotiation with restructuring typically results from a bargaining game between firm claimants (Fan and Sundaresan 2000). Because bargaining with restructuring such as debt-equity swaps is reported to be prohibitively costly and time consuming in reality, I abstract away from restructuring. Anderson and Sundaresan (1996) investigate restructuring, but assign all the bargaining power to equityholders. However, if debtholders refuse the proposal of the exchange offer at renegotiation, equityholders will be much worse off. Hence, as long as debtholders have any bargaining power, the outcome of renegotiation with restructuring will leave debtholders above their indifference point. In my model, the participation constraint for debtholders at renegotiation is not binding for most parameter specifications.

Condition (10) is satisfied for reasonable parameter values if the payout ratio is not too high. As dividend restrictions represent the most frequently used covenants in both private and public debt (Bradley and Roberts 2004), this condition is likely to be satisfied in reality.

The continuation value of equity is given by the difference between the value of the firm and the value of debt:

$$\begin{aligned}E(V,L) = {} & V - \frac{\delta V_B}{\lambda^2+(\sigma+b)\lambda}(\frac{V}{V_B})^{-\xi} - \delta\frac{V_B}{\lambda}(\frac{V}{V_B})^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma-b} \\ & + \frac{(\tau-1)c+\Gamma(L)}{r}\big(1-(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{V}{V_B})^{-\xi}\big) \\ & + (\delta-\rho)\big((1/\delta)\big(V - \frac{\phi(-(\sigma+b)\sqrt{d})}{\phi(\lambda\sqrt{d})}V_B(\frac{V}{V_B})^{-\xi}\big) \\ & - \big(V/\delta - \frac{V_B}{\lambda^2+(\sigma+b)\lambda}(\frac{V}{V_B})^{-\xi} - \frac{V_B}{\lambda}(\frac{V}{V_B})^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma-b}\big)\big)\end{aligned} \tag{11}$$

In the presented setting with corporate cash, shareholders hold a Parisian down-and-out call option on the firm's assets. That is, they have a residual claim on the cash flows generated by the firm unless the value of the assets reaches the distress threshold and remains below this threshold until the (pledged) cash is exhausted. The value of equity at initiation is equal to the continuation value of equity, minus the initial cash (L) contributed by equityholders. Note that a cash policy of $L = 0$ induces $d = 0$, and $\Phi(0) = 1$. In this special case, Equation (11) simply collapses to the equity value in Leland (1994), i.e., to $V + \frac{(\tau-1)c}{r}(1-(\frac{V}{V_B})^{-\xi}) - V_B \frac{V}{V_B}^{-\xi}$, where the liquidation time is given by the first passage time of the state variable to V_B. When L goes to infinity, economic distress never leads to liquidation. Debt essentially becomes riskless but $\Gamma(L)$, and, hence, the equity value tend towards minus infinity.

The timing of economic distress depends on the threshold selected by shareholders. For each coupon level, this threshold can be characterized by the economic distress policy which maximizes the equity value. It satisfies the smooth-pasting condition (Dumas 1991). Hence, the economic distress threshold where equityholders optimally stop contributing funds is given by

$$\frac{\partial E(V,L)}{\partial V}|_{V=V_B^*} \stackrel{!}{=} 0, \tag{12}$$

which yields

$$V_B^* = \frac{\frac{\Gamma+(\tau-1)c}{r}\big(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\big)\xi}{(\rho-\delta)\frac{1}{\delta}\frac{\Phi(-(b+\sigma)\sqrt{d})}{\Phi(\lambda\sqrt{d})}(1+\xi)-\rho\big(\frac{1}{\lambda^2+(\sigma+b)\lambda}+\frac{1/(\lambda-\sigma-b)}{\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\big)(1+\xi)}. \tag{13}$$

3.2. The target cash policy

The total private benefit a manager receives from a firm is given by
$\mathbb{E}_Q\big(\int_0^{\theta V_B} e^{-ru}\omega du\big)+\Psi(E(V,L)-L)+\mathbb{E}_Q\big(\sum_{i=1}^{N_{\theta V_B}} e^{-rt_i}\eta log(L)1_{V_u>V_B}\big)$.
It can be expressed in closed form:

Proposition 3. *Managers' total private benefit is given by*

$$\begin{aligned} U(V,L) &= \frac{\omega}{r}\big(1-\frac{e^{a\sqrt{2r}}}{\phi(\sqrt{2rd})}\big)+\Psi(E(V,L)-L) \\ &+\frac{y\eta log(L)}{r}\big(1-\big(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\big)\big(\frac{V}{V_B}\big)^{-\xi}\big), \end{aligned} \tag{14}$$

with $a=(1/\sigma)(log(V_B^*/V))$, *and* $b=(1/\sigma)(r-\delta-\sigma^2/2)$.

Proof. See in the Appendix ☐

The first term in Equation (14) is the value of the fixed income ω obtained up to default. The second term expresses the value of the equity based compensation Ψ. The last term, finally, corresponds to managers' personal utility from investing (spending) cash.

The target cash policy L^* is chosen to maximize managers' private benefit subject to two constraints. The first requires that equityholders are initially willing to contribute the demanded cash (participation constraint).[10] The second is that the incumbent management can preclude a takeover because the firm value is larger than the firm value F^A under the alternative management team minus transaction costs.[11]

$$\begin{aligned} L^* &:= \text{argmax}_L U(V,L) \\ s.t. \quad & F(V,L)\geq F(V,0)+L \\ & F(V,L)\geq F^A(V,L^A)-T. \end{aligned} \tag{15}$$

[10] This condition also ensures that $R(L^*)>L^*$, because holding cash is costly.
[11] The second condition can also be interpreted in alternative ways: For example, T may represent costs of bargaining between managers and equityholders, and F^A the firm value with the expected bargained level of cash holdings.

The System (15) shows that within certain bounds, managers have full control over the choice of the corporate cash policy L^*. This control has two direct implications. First, it is optimal for managers to capture the investment opportunities when they occur as long as the firm is not in economic distress. The reason is that investing increases each term in Equation (14). Second, managers' benefit is zero at default. Hence, they optimally use (pledge) cash to delay default as soon as equityholders stop injecting funds because their total benefit at V_B^* from continued operation, i.e., $U(V_B^*, L^*)$, is always larger than 0.

3.3. Optimal debt policy

While the distress policy is chosen (ex-post) by equityholders to maximize the value of equity after the issuance of corporate debt and management's choice of the cash policy, the optimal debt policy maximizes $E(V, L^*) - L$ plus the proceeds from the debt issue.[12] As $F(V, L^*)_L = E(V, L^*) - L + D(V, L^*)$, the optimal coupon is chosen by equityholders to maximize the value of the firm, i.e., $c^* := argmax_c F(V, L^*) - L$.

4. Empirical implications

This section discusses the results and empirical implications of the model for credit risk, financial policy, excess cash holdings, and agency costs.

4.1. Parameter choice

Input parameters are set in accordance with the literature to reflect a typical S&P 500 firm (François and Morellec 2004, Hackbarth, Miao, and Morellec 2006): The initial value of the firm's invested assets, $V = 100$; the riskless interest rate, $r = 6\%$; the net tax advantage of debt, $\tau = 0.15$; liquidation costs, $(1 - \alpha) = 40\%$; costs of economic distress, $\rho = 3\%$; the constant payout rate, $\delta = 5\%$; the volatility of assets in place, $\sigma = 0.2$. Following the quotes in Meeks and Meeks (2001), takeover costs are assumed to be 5% of the unlevered asset value. Additionally, I need to specify the compensation of managers, the yield of cash holdings, and the profitability of investments. In the baseline scenario, I choose ω and Ψ to reflect that equity based management compensation accounts for about 60% of the total monetary compensation for S&P 500 firms (Bebchuk and Grinstein 2005).

[12] Managers need to communicate the cash policy associated with each debt policy at initiation, because they must raise the necessary cash from equityholders. Given the knowledge of the cash policy, debtholders anticipate the threshold V_B^* for each coupon level.

I start without the personal utility from investing cash as this term is not observable in reality. As an extension, I also incorporate managers' private utility from investing by setting η such that the expected utility from investment represents 10%, or 20% of their overall utility. The yield of very liquid assets (cash) is set to $l = 2\%$.

For the curvature of the profit function of invested capital, Riddick and Whited (2009) obtain a γ around 0.75, and Hennessy and Whited (2007) one around 0.63. I choose $\gamma = 0.7$ in the profit function of investment opportunities (AL^{γ}). Anderson and Reeb (2003) and Loughran and Ritter (1997) find an average operating income (EBITDA) to assets ratio around 15% for Compustat and S&P 500 firms, respectively. I calibrate the parameter A such that the immediate NPV of investing L^* in the opportunity corresponds to the (after tax) NPV of the same amount invested in assets yielding a 15% operating income. This calibration yields $A = 4.2$. I also validate my results with alternative choices for A. Finally, the intensity of the occurrence of investment opportunities is set to $y = 0.1$, such that the model's predicted cash ratio equals 11.04 for the average firm. This ratio corresponds to the average cash ratio observed for S&P 500 firms between 1999 and 2009 outside the information technology sector. In the baseline specification, I assume that $y = y^A$. While the quantitative results vary with these parameters, I show that the important predictions do not critically depend on their choice.

4.2. Optimal distress policy of shareholders

Once the cash policy is determined and debt has been issued, shareholders' only choice is the ex-post selection of the economic distress policy. In particular, they choose the distress threshold which maximizes the value of equity, that is, the optimal asset value where they stop injecting funds into the firm. Default, however, is triggered when managers are unable to continue service coupon payments during economic distress by pledging cash, i.e., when cash is exhausted.

When choosing the optimal economic distress policy in a firm with cash holdings, equityholders trade off their benefit of economic distress, i.e., sparing (after tax) debt service without losing potential future income, against distress costs (δV), the foregone total expected instantaneous yield from cash holdings ($\Gamma(L)$) generated by the yield of cash (l) plus the expected payouts from investment opportunities, and the probability of bankruptcy. Figure 1 depicts the optimal distress threshold of a firm with cash holdings of $L/V = 0.1$ (solid line). The dashed line shows the distress policy analog of a firm with identical total payoff, but without cash holdings. Note that in standard structural models without cash, the distress threshold corresponds to the default threshold, as firms immediately default when equityholders stop contributing funds.

INSERT FIGURE 1 HERE

Figure 1 reveals that equityholders stop contributing funds earlier in a firm with cash than in an otherwise identical firm without cash.[13] The intuition for this important result is as follows: In standard structural default models, equityholders keep servicing debt until the distress threshold is reached, even if the net-worth of a firm is already negative. The reason is that, while protected on the downside by limited liability of equity, they can still profit from a potential rebound of the firm value. Ceasing to inject funds triggers immediate default, which results in a loss of the potential upside. With cash holdings, however, equityholders anticipate that the firm will not immediately default when they stop injecting funds. In case of substantial negative net-worth, they can consequently cease to finance the debt service while still observing how the firm evolves thereafter. The possibility to save coupon payments during economic distress while maintaining the upside if the firm rebounds explains why equityholders stop injecting funds earlier in a firm with cash holdings.

The finding in Figure 1 challenges the traditional view that cash holdings simply reduce default risk. I show that the total impact of cash on credit risk can be divided into a direct, and an indirect effect. The obvious, direct effect is that higher cash admits managers more time to defer default when shareholders stop contributing funds, which decreases credit spreads. The indirect effect occurs because equityholders stop contributing funds earlier for larger cash holdings. Hence, higher cash makes it harder for firms to raise external funds when they perform bad.

Table I explores the quantitative impact of cash on credit risk. The second column shows the direct marginal effect of cash on credit risk. It corresponds to the marginal decrease of a firm's credit spread if one additional unit of cash is added to the optimal level, given that the default boundary is left constant at the one of an otherwise identical firm with optimal cash holdings. The third column shows the indirect marginal effect caused by the fact that equityholders optimally increase the default threshold with an additional unit of cash. It is given by the credit spread of a firm with $L^* + 1$ units of cash (and with an optimal default threshold for this cash level) minus the credit spread of the otherwise identical firm with a hypothetical default threshold equal to the one of a firm with optimal cash holdings L^*. Finally, the total marginal effect of cash on credit risk is

[13] When the debt service is lower than the expected hypothetical payout from cash in the firm without cash, equityholders will, in fact, never stop contributing funds in this example (see the solid line below 1.1).

captured in the fourth column as the total decrease in the credit spread if one additional unit of cash is added to a firm with optimal cash holdings. In the baseline scenario, the credit spread is 43 basis points (bps). The direct marginal effect of cash reduces credit spreads by 2.8 bps. The indirect marginal effect, however, increases credit spreads by 1.63 bps, canceling a large portion (58%) of the direct marginal effect. As a result, the total marginal reduction in credit spreads from additional cash is only 1.17 bps. The importance of the indirect effect could explain why, despite the intuitive appeal and widespread use of liquidity proxies in empirical default-predicting models, many studies conclude that liquidity measures are not significantly negatively associated with default (Zmijewski 1984, Begley and Watts 1996, Shumway 2001, Hillegeist, Keatin, Cram, and Lundstedt 2004).

In the last column, Table I also shows the total marginal effect of cash on the value of a fixed income stream of 1 obtained up to default.

INSERT TABLE I HERE

The results in Table I suggest that the effect of cash holdings on credit risk depends on firm characteristics. During economic distress, equityholders forego the instantaneous yield of cash because cash is pledged, incur distress costs, and lose the tax shield. Hence, when l, ρ, and τ are lower, equityholders' opportunity costs of economic distress decrease which enhances their incentive to stop contributing funds early. Graphically, the dashed line in Figure 1 moves to the upper left. As a result, the indirect marginal effect becomes stronger compared to the baseline scenario which causes a weaker total reduction of credit risk from additional cash.

Next, I investigate how the results are related to the payout ratio. Decreasing the payout ratio raises the risk-neutral drift of the asset value which increases the probability of recovery from economic distress. One unit of additional cash, consequently, enhances the value of the fixed income stream more than in the baseline scenario. Table I, however, also shows that the total marginal effect of cash on credit risk decreases because debtholders lose coupon payments in case of recovery from economic distress.

A higher asset volatility particularly increases the direct marginal effect of cash on credit risk, because the probability of reaching a given economic distress threshold is larger. As a consequence, the total marginal effect of cash becomes stronger than in the baseline parameter setting.

In sum, the model generates specific empirical predictions, namely that the marginal impact of cash on credit spreads should be weaker for a lower yield of cash, lower economic distress costs, a lower tax shield, a smaller payout ratio, and a lower volatility of assets.

4.3. Corporate debt policy choice

When applied to financing decisions, standard structural models typically generate leverage ratios that exceed those observed in practice. This observation is called the under-leverage puzzle. I show that one potential explanation for this limitation is that standard models overlook important determinants of the corporate debt policy, namely investment opportunities and cash holdings.

Equation (5) indicates that debt financing affects firm value in several ways. First, it provides a tax shield. Second, debt affects distress and bankruptcy costs by changing the corresponding distress and bankruptcy policy. Third, debt causes underinvestment as cash is pledged and not invested during economic distress, and it also induces a loss of the investment opportunities upon default.

Without cash holdings, the optimal leverage corresponds to 50.3% for my baseline parameters. If investment opportunities and cash holdings are incorporated, equityholders choose a significantly lower leverage of 42.5% when they trade off the costs and benefits of leverage. There are several reasons for this result. First, cash makes debt more expensive because equityholders optimally stop contributing funds earlier than in a firm without cash. Early economic distress is costly because the yield of cash is lost, economic distress costs accrue, and the tax shield is foregone when the firm is unable to raise external equity. Second, cash affects the costs of debt via the default timing. It allows managers to deviate from a default policy which maximizes the ex-post value of equity.[14] Third, the recovery rate of total assets decreases when incorporating investments, because these opportunities are lost in case of default. Finally, investment opportunities and the associated profits are foregone during economic distress. The underinvestment problem of debt is, consequently, more severe with larger cash holdings as equityholders stop contributing funds earlier.

4.4. Corporate cash policy and excess cash

I now turn to understanding the extent to which managerial control over cash influences the corporate cash policy. Managers trade off their private benefits and costs of cash according to Equation (15) when determining the cash policy L^*. The discretionary use of cash allows them to avoid immediate default in case of economic distress, and, therefore, to extend the period they obtain the

[14] The direction of this effect on the ex-ante value of equity, however, depends on the leverage ratio because the deviation also influences the issue price of debt.

fixed income (and utility from investments). This function of cash as a buffer against bankruptcy in hard times induces managers to target excess cash, i.e., a cash level beyond the target L^{SH} of equityholders (self-preservation motive).

INSERT TABLE II HERE

Table II describes the corporate cash policy choice and its determinants in an unrestricted firm without a market for corporate control. Besides the shareholder wealth maximizing cash ratio (L^{SH}/V) and the corporate cash ratio (L^*/V), the table also reports excess cash ratios ($(L^* - L^{SH})/V$). In the baseline parameter specification, the cash ratio (L^*/V) targeted by managers corresponds to 11.04%, which is considerably higher than the equity value maximizing cash level of 9.22%.[15]

Next, Table II investigates the impact of management compensation on the cash policy. The results are very sensitive to the relation between fixed and variable compensation. The lower the equity part Ψ, the higher managers' incentive to hold excess cash. The reason is that the equity share in their compensation package internalizes the costs of deviating from a shareholder value maximizing strategy. In fact, managers target a cash and excess cash ratio of 26.46% and 17.09%, respectively, when they are only compensated with a fixed wage. The takeover constraint in the optimization problem (15) is binding in this case, as managers wish to raise an infinite amount of cash when they set up the firm. The findings confirm the empirical results in Marchica and Mura (2008) who show that firms with persistent high cash holdings seem to have significantly lower managerial ownership. When the equity part accounts for 30% of the overall monetary compensation, the target cash ratio is 14.20% which is still 4.73% of assets larger than the shareholder wealth maximizing level. As expected, excess cash vanishes when managers are almost exclusively (90%) compensated on an equity-linked basis.[16] Finally, the table also shows that excess cash is strongly positively related to managers' private utility from investing cash η. One could easily extend the present analysis to include a private utility for managers from retaining control which would simply induce a higher portion of the fixed compensation, and, consequently, larger excess cash.

The analysis of the yield on very liquid, cash-like securities l indicates that, as expected, the shareholder wealth maximizing cash ratio decreases with lower l because the opportunity costs of

[15] I also calibrate the parameter A of the profit function to an average operating income to assets ratio of 12% and 18%, respectively. The corresponding excess cash ratios are 1.01% and 2.97%, respectively.

[16] Of course, equity-linked compensation induces other costs not captured in my model, see for example Hall and Murphy (2002).

holding liquid assets, given by $r - l$, increase. To explain the amount of excess cash, I need to investigate mangers' trade-off when choosing the target cash policy. They balance the positive effect of excess cash on the value of their fixed income against the negative effect on their equity compensation. The previous Table I shows that the former is weaker for lower l. At the same time, excess cash has a stronger negative effect on the value of the equity compensation due to the higher opportunity costs of holding cash. Both effects attenuate managers' propensity for excess cash when l becomes lower. These results are in line with Kim, Mauer, and Sherman (1998) who find empirically that both corporate cash holdings and excess cash decrease with the opportunity cost of holding liquid assets.

Table I indicates that the self-preservation motive of managers is weaker for smaller distress costs ρ and a lower tax advantage of debt τ. At the same time, the impact of excess cash on the equity part of the compensation is only moderately affected by these parameters. Hence, managers target lower excess cash when ρ or τ decrease. Finally, the last two rows in Table II show that a lower payout ratio is associated with larger excess cash, mainly because the total marginal effect of additional cash on the value of managers' fixed income increases for lower δ.

The explanations to the results in Table II illustrate that to assess managers' propensity for excess cash due to the self-preservation motive, it is essential to understand how cash holdings affect credit risk.

Cash is costly to equityholders because of its holding costs and the fact that it accrues to debtholders upon default. For excess cash, these costs do not outweigh the benefits from the additionally admitted investments. To estimate the value loss to equityholders from the misalignment of incentives, I compute ex-post agency costs. They are calculated as the percentage reduction in the ex-post equity value of a firm with cash holdings equal to L^* compared to an otherwise identical firm with holdings equal to L^{SH}.[17] This exercise is useful because it assists in the interpretation of the economic magnitude of the impact of excess cash on the equity value of firms.

In the baseline setting, agency costs are relatively moderate at 0.17% of the equity value. The magnitude of equityholders' loss is within the tolerance on the part of shareholders for value destruction, in which the limit is imposed by any costs of assuming control of the corporation.

[17] Managers' cash policy is an ex-post financial policy. They maximize the valuation of their claims after debt has been issued. Debtholders anticipate the cash policy choice when debt is issued. The ex-ante agency costs to equityholders of levered firms, given by the percentage reduction in the ex-ante firm value with L^* compared to the one with L^{SH}, are lower than the ex-post agency costs. In fact, the firm value even increases under L^* for many parameter combinations. The reason is that excess cash makes debt safer, which induces higher debt issue proceeds. Hence, even without takeover costs (or negotiation frictions), equityholders may, ex-ante, agree that managers control the cash policy decision.

Agency costs are, however, highly dependent on the parameter specification. For example, with 30% equity compensation they increase to 0.91%, and with 20% utility from investing even to 5.66%.

Table II shows that excess cash holdings can be large in certain firms, especially when, as in reality, compensation schemes can not be solely designed to implement the optimal cash policy. This finding motivates to analyze to what extent the threat of a control challenge prompts managers to target a more efficient cash level. The impact of the market for corporate control depends on the opportunity costs of replacing the management faced by investors. The opportunity costs are determined by the ability of the alternative management team, captured by the intensity y^A, and by the takeover costs T. Table III summarizes the results for a firm with managers that obtain 60% of their monetary compensation from equity, and 10% of their total utility from investing cash. Panel A indicates that the higher the costs T imposed by control challenges, the larger are the selected corporate cash and excess cash ratios. In the last row, takeover treats are too weak to influence the management's choice of the target cash policy. Panel B additionally investigates the impact of the alternative management's ability to find profitable investment opportunities, y^A, on corporate cash holdings. The resulting corporate cash ratios and excess cash ratios are larger if the ability of the alternative management is lower than the one of the incumbent. Overall, the results suggest that because of their specific human capital, and the costs imposed by control transactions, self-interested managers are partially entrenched which allows them to exploit the reduced threat of takeovers by targeting larger excess cash.

INSERT TABLE III HERE

The outcome in Table III reflects the results in the empirical literature. Dittmar, Mahrt-Smith, and Servaes (2003), for example, confirm that entrenched managers are more likely to build excess cash balances, and that greater shareholder rights (captured by smaller takeover costs in my model) are associated with lower excess cash. Pinkowitz, Stulz, and Williamson (2003) show that firms in countries with poor investor protection hold more excess cash, and that liquid assets in those countries contribute substantially less to firm value than the ones held by firms in countries with stronger investor protection. In my model, cash is indeed less valuable to firms when takeover costs are larger due to the higher excess holdings. The results in Kalcheva and Lins (2007) propose that value discounts apply to firms expected to be exposed to relatively severe managerial agency problems if they have a combination of managerial entrenchment, high cash holdings, and poor external protection against expropriation. Consistent with this idea, my model assigns the highest agency costs (discount) when managerial entrenchment is combined with high takeover costs. For example, agency costs are 0.61% of the equity value for $T = 0.1$ and $y^A = 0.1$, but correspond to 2.26% for $T = 0.9$ and $y^A = 0.098$. Similar thoughts are reported by Yun (2009) who shows that firms increase cash holdings relative to lines of credit when state antitakeover laws reduce future control threats and when internal governance is weak. Harford, Mansi, and Maxwell (2008) argue that more managerial entrenchment may increase the transaction costs bounds within which managers can operate before it becomes worthwhile to remove them, but they still must be careful not to accumulate large, unused cash stockpiles. Faleye (2004) confirms this argument by showing that proxy contests are increasing in excess cash reserves. Moreover, executive turnover and cash distributions to shareholders rise, while cash holdings significantly decline following such contests. Overall, the empirical evidence is consistent with the idea that self-interested managers balance the private benefits of cash against increased takeover threats.

In reality, shareholders can also rely on alternative management control mechanisms to induce managers to deviate from their personal target cash level towards the one which maximizes the value of equity. The presented analysis could be analogously applied to such mechanisms, or to a costly bargaining game between shareholders and managers on the cash policy.

4.5. Do riskier firms hold a higher amount of cash?

In this section, I derive the relationship between asset risk and corporate cash holdings. I first explain the relationship between the equity value maximizing cash ratio and the asset volatility. Then, I discuss how managers' target excess cash depends on the riskiness of assets. Finally, I show that when managers' propensity for excess cash dominates the relationship between the equity value maximizing cash ratio and the asset volatility, the relationship between corporate cash holdings and asset volatility turns positive.

Recent empirical studies document that riskier firms tend to hold more cash (Kim, Mauer, and Sherman 1998, Harford, Mansi, and Maxwell 2008, Iskandar-Datta and Jia 2010, Mahmudi and Pavlin 2010, Acharya, Davydenko, and Strebulaev 2011). The standard reasoning is that cash holdings are valuable to constrained firms because they allow to avoid passing up valuable investment opportunities when other sources of funds are unavailable. Firms in industries with more volatile cash flows are then expected to hold more liquidity because of the higher probability of cash flow shortfalls.[18] Such a rational from a precautionary motive does not maintain once I incorporate agency conflicts. Figure 2 shows the relationship between the cash ratio which maximizes shareholder wealth (L^{SH}/V) and the asset volatility.

INSERT FIGURE 2 HERE

For a wide range of asset volatilities, the relationship is negative. The reason is that the expected return from cash to equityholders decreases with credit risk because managers use cash to defer default instead of investing during economic distress (see Section 2), and because cash accrues to debtholders upon default.[19]

Figure 3 plots excess cash against the asset volatility, and shows that they are positively related.

INSERT FIGURE 3 HERE

The intuition for this result directly evolves from managers' control over cash holdings. According to their trade-off described in Equation (14), managers have a propensity towards excess cash to increase the value of their fixed income. Table I shows that when the asset volatility increases, this self-preservation motive becomes stronger. Hence, managers of risky firms will optimally target higher levels of excess cash.

Finally, as the target corporate cash ratio L^*/V corresponds to the sum of the shareholder wealth maximizing cash ratio and excess cash, the relationship between asset risk and L^*/V depends on whether the effect in Figure 2 or 3 dominates. Figure 4 shows that the latter dominates in the baseline parameter specification such that the relationship between asset risk and the corporate cash ratio (L^*/V) turns positive.

INSERT FIGURE 4 HERE

This result suggests that agency conflicts, i.e., managers endogenously targeting higher excess

[18] Acharya, Davydenko, and Strebulaev (2011) argue that in the presence of restrictions on external financing, and when future cash flows are only partially pledgable, this result is driven by firms' trade-off between investing available cash in a profitable long-term project and retaining cash in order to reduce the likelihood of losing the future cash flows from the project due to default. The incentive of equityholders to inject cash when liquidity is scarce, however, increases with the long-term expected cash flows from previous investments in profitable projects. Therefore, incorporating endogenous external financing clearly dampens the described trade-off. Moreover, one hardly observes firms with plenty of profitable long-term projects defaulting due to a temporary liquidity shortfall (Davydenko and Strebulaev 2007).

[19] For very high asset volatilities, the credit risk decreases with the volatility. The reason is that riskier firms optimally initiate a lower coupon. A lower coupon allows a longer survival time during economic distress for a given amount of cash. At high levels of volatility, this effect dominates the one of the asset volatility on the probability of reaching distress.

cash in riskier firms, are responsible for the empirical finding that cash holdings are positively associated with credit risk. Mauer and Liu (2011) provide striking empirical evidence for this view. The authors show that firms which encourage CEO risk-taking hold more cash. Importantly, they also find that the marginal value of cash to equityholders decreases for firms with high risk-taking incentives which supports my hypothesis that the positive relation between cash holdings and CEO risk-taking is driven by a desire on part of managers to hold excess liquidity.

Bates, Kahle, and Stulz (2009) argue that the recent increase in corporate cash holdings is concentrated among small firms in industries that experienced the greatest increase in idiosyncratic cash flow volatility. As Himmelberg, Hubbard, and Palia (1999), and Baker and Hall (2004) find that CEOs of smaller firms have a lower dollar incentive from equity compensation, and because a higher portion of fixed income even enforces the positive relationship between managers' target excess cash and asset risk, my model suggests that agency is, at least partially, responsible for the reported pronounced increase of corporate cash holdings within smaller firms which became riskier.

A normative implication from this insight is that riskier firms should implement stronger mechanisms to limit excess cash holdings than firms with lower risk. As long as these mechanisms are costly, one would, however, still observe the proposed relationship.

4.6. Simulation approach

In this section, I adapt a simulation approach in the spirit of Strebulaev (2007), and Bhamra, Kuehn, and Strebulaev (2010) to explore the empirical cash properties of a model-implied economy of firms which is structurally similar to the real economy. It is shown that, once the intensity y is calibrated to observed corporate cash levels, the self-preservation motive predicts an amount of excess cash which induces cash-properties in the model-implied economy comparable to their empirical counterparts.

Following a simulation approach is important for several reasons. First, the previous results are obtained under the assumption that firms are at issue. In reality, however, they are mostly not at issue. Over time, firms deviate from the initially optimal structure because they do not continuously adapt their cash and debt policies. Only a simulation approach can explore the impact of these deviations on observed cash properties in an economy.

Second, as Bhamra, Kuehn, and Strebulaev (2010) argue, it is not clear that the results of a structural model for a typical individual firm can be compared to the empirical average results for a sample of firms of which the typical firm is representative. The reason is that real firms in a sample differ widely in their firm characteristics, and that parameters of interest are non-linear in these characteristics. Therefore, it can be crucial to generate a sample implied by the model which has similar characteristics than the real firm sample, and to compare the resulting average parameter of interest to its empirical counterpart.

Third, the main feature modeled by this paper is that cash is used by managers to service debt during economic distress. Cash, therefore, mainly accrues to debtholders when firms approach default. At initiation, however, firms are not close to default because they are structured optimally. Hence, to investigate the model-implied cash properties of firms approaching default, and to compare them to the ones of real firms in an economy, a simulation approach is necessary.

In order to obtain the empirical implications of my model, I imitate the true dynamics of S&P 500 firms by simulating over time a cross-section of firms that resembles its empirical counterpart. I start with a range of different firms which are initially structured optimally. In particular, I consider firms with asset volatilities from 0.15 to 0.25 (step size 0.025), with intensities of the occurrence of investment opportunities from 0.025 to 0.3 (step size 0.025), and with managers that have an η, i.e., a private utility from investing cash, from 0 to 4 (step size 1). Parameter combinations which do not satisfy the participation constraint of equityholders, and the takeover constraint in System (15) are deleted. Each of the remaining 257 parameter combinations is quarterly simulated forward 100 times over 10 years, which leaves me with a model generated economy of 25700 different firms.

Next, for every year over the period from 1999 to 2009, I take the empirical sample of S&P 500 firms from Compustat, and construct the average cross-sectional distribution of their quasi-market leverage and cash ratio. The quasi-market leverage is defined as the ratio of book debt to the sum of book debt and the market value of equity. Book debt is total assets (item 6) minus book equity (Baker and Wurgler 2002). Book equity is total assets minus total liabilities (item 181) minus preferred stock (item 10, replaced by item 56 when missing) plus deferred taxes (item 35) plus convertible debt (item 79). The cash ratio corresponds to cash and marketable securities (item 1) divided by total assets (item 6). The information technology sector is excluded, because it is characterized by firms with very low leverage and very large cash holdings often exceeding their entire asset value. It is difficult to replicate such firms with reasonable parameter values without violating the participation constraint in System (14). An alternative model of cash holdings tailored to this sector would be needed. My results are only marginally affected by the exclusion of the financial and utility sector.

I then match the average historical distribution of S&P 500 firms with its model-implied counterpart in the generated economy. In particular, for each observation in the empirical sample, I select the observation in the model generated economy with the minimum distance regarding the deviation from the empirical average quasi-market leverage and cash ratio. The matching is generally very accurate, with a median Euclidean distance of 0.0032.[20] That is, the median firm is matched with the root of the sum of the squared deviations being 0.0032. This matching procedure allows me to construct a cross-sectional distribution of model-implied firms which is almost identical to the one empirically observed. Note that I do not assume that the initial model generated economy consists of firms which are similar to S&P 500 firms. Rather, I choose only those firms within this economy which can be identified as being very similar to S&P 500 firms.

Finally, I simulate the matched model-implied firms forward for 10 years with quarterly time intervals, which gives me a simulated data set. Table IV presents the properties of cash within this simulated data set. They are then directly compared to empirical studies. I concentrate on the main cash features considered important to be explained by a model of corporate cash holdings such as observed levels, market values, and the marginal values of cash. To explore the distributional features of my results, I repeat the entire simulation 1000 times, and report the average, the mean, and the 25^{th} and 75^{th} percentiles.

INSERT TABLE IV HERE

The average cash ratio in my simulated data set is 11.37, very close to the average of 11.04 in the empirical sample of S&P 500 firms. This result is not surprising, as the cash ratio is matched in the matching procedure.

In Panel A of Table IV, I report the average as well as distributional properties of the ratio of the market value of cash to its face value for my simulated sample of firms. The market value of cash is the total value to the firm of holding cash. The face value simply corresponds to the nominal amount of cash. The average ratio is 1.2959 which reflects the corresponding estimate of 1.25 from Pinkowitz and Williamson (2002) very well.

Faulkender and Petersen (2006) report the marginal value of cash for an average Compustat firm over the 1971 to 2001 period. Their estimate of the impact of a change in cash holdings suggests that an extra dollar of cash is, on average, only valued by shareholders at 0.75. When they allow the change in cash to interact with the level of cash and with leverage, the estimated

[20] The average Euclidean distance is higher at 0.033 due to a few firms with very low leverage, but very high cash ratios which my model matches with a relatively large Euclidean distance.

marginal value of cash is 0.94 at the mean leverage and mean cash ratio. Dittmar and Mahrt-Smith (2007) derive similar values for poorly governed firms, and Pinkowitz, Stulz, and Williamson (2006) obtain comparable results with international data. The authors also show that the marginal value of cash is above 1 for low levels of cash, but decreasing with the firms' cash positions to levels below 1. This finding is consistent with the idea that there is an upper bound on the amount of cash for which a firm is rewarded, but that the mean firm holds cash in excess of that upper bound due to the self-preservation motive of managers. Panel B investigates the marginal value of cash to equityholders in my simulated data set. It is calculated numerically as the first derivative of the continuation value of equity in Equation (11) with respect to L at model-implied corporate cash levels. The resulting average value in the simulated data set is 0.6938 which is very close to the marginal value of cash reported in Faulkender and Petersen (2006).[21]

The results show that the model, once calibrated to empirically observed corporate cash levels, is able to replicate the observed market value of cash to firms, as well as the marginal value of cash. Hence, the self-preservation motive captures the empirically reported properties of cash quite well

Panel C of Table IV shows the mean, the median, and the 25^{th} and 75^{th} percentiles of the regression coefficient of leverage on the marginal value of cash in my simulated data set. A simple OLS regression with firm fixed effects is applied. Faulkender and Petersen (2006) argue that if firms use cash to pay down debt, a small increase in cash reserves should partially go to increasing debt value, not solely to increasing equity value. Thus, the marginal value of cash to equityholders should be larger for firms with low leverage than for those with high leverage. The authors show that the marginal value of cash to equityholders is, indeed, decreasing with leverage. As my model captures that cash accrues to debtholders in case of economic distress, the coefficient of leverage in the simulated data set is also negative, and very stable around a mean of -0.1176. This result suggests that managers' control over cash is a realistic feature.

Finally, as the focus of this paper is on the self-preservation motive driven by the impact of cash on default, it is important that the presented model is able to quantitatively explain the empirical determinants of credit risk. Table V, therefore, lists the coefficients of a OLS-regression of credit spreads on the cash ratio, leverage, and asset volatility in my simulated data set. The coefficients on leverage and asset volatility are quantitatively very similar to their empirical counterparts reported in Acharya, Davydenko, and Strebulaev (2011). While the coefficient on the cash ratio has the

[21] For the benchmark firm with optimal leverage, the marginal value of cash corresponds to $\frac{\partial E(V,L^*)}{\partial L} = 0.79$ at initiation. It is slightly higher than the average in the simulation mainly because the marginal value of cash decreases when firms approach default, but does hardly increase when X rises above its initial value.

expected sign, its magnitude is not directly comparable to the one of the coefficient in their Table 5 because the authors apply an instrumental variable approach.[22]

INSERT TABLE V HERE

5. Extension

In this section, I consider the impact of correlated investment opportunities on corporate cash holdings.

5.1. Corporate cash policy and correlated investment opportunities

In reality, investment opportunities are likely to be correlated to the value of assets. To capture this feature in a simple fashion, I assume that the Poisson intensity of the occurrence of investment opportunities follows

$$dy_t = \sigma_y y_t dW_t^y, \tag{16}$$

with $E(dW_t dW_t^y) = \beta dt$. When y_t is stochastic, the distress policy of equityholders can not be expressed in term of a fixed asset value (V_B), because their decision to stop injecting funds depends on the pair (V_t, y_t). To reduce the number of state variables from 2 to 1, I define $A_t = V_t + E_Q \int_t^\infty (R(L) - L) y_u e^{-ru} du = V_t + \frac{(R(L)-L)}{r} y_t$ which comprises the entire stochastic part of the total asset value. This definition of A_t allows to characterize the distress policy in terms of a constant A_B. To simplify calculations, it is assumed that $\delta = r$, and that debtholders also recover the fraction α of the value of the investment opportunities upon default. Proposition 4 expresses the value of equity and debt with correlated investment opportunities.

Proposition 4. *If the Poisson intensity of the occurrence of investment opportunities is correlated*

[22] In their plain OLS regression in Table 4, Acharya, Davydenko, and Strebulaev (2011) obtain a positive coefficient for the cash ratio. I also get a positive coefficient of similar magnitude when regressing credit spreads on the cash ratio without controls. This result, however, occurs because the unobserved X drives both credit spreads and the cash ratio.

to the value of assets, the continuation value of equity for a given default policy (A_B) is given by

$$
\begin{aligned}
E(V,L,y) = A &- \frac{rA_B}{\lambda^2+(\sigma_A+b)\lambda}(\frac{A}{A_B})^{-\xi} - r\frac{A_B}{\lambda}(\frac{A}{A_B})^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma_A-b} \\
&+\frac{(\tau-1)c+lL}{r}\big(1-(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{A}{A_B})^{-\xi}\big) \\
&+(r-\rho)\big((1/r)\big(A-\frac{\phi(-(\sigma_A+b)\sqrt{d})}{\phi(\lambda\sqrt{d})}A_B(\frac{A}{A_B})^{-\xi}\big) \\
&-\big(A/r-\frac{A_B}{\lambda^2+(\sigma_A+b)\lambda}(\frac{A}{A_B})^{-\xi}-\frac{V_B}{\lambda}(\frac{V}{V_B})^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma_A-b}\big)\big),
\end{aligned}
\tag{17}
$$

and the value of debt corresponds to

$$
D(V,L,y) = \frac{c}{r}\big(1-(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{A}{A_B})^{-\xi}\big)+\alpha A_B(\frac{A}{A_B})^{-\xi}\frac{\Phi(-(b+\sigma_A)\sqrt{d})}{\Phi(\lambda\sqrt{d})}, \tag{18}
$$

where $\Phi(x) = 1 + x\sqrt{2\pi}exp(\frac{x^2}{2})N(x)$, N is the standard normal cumulative distribution function, $\sigma_A^2 = (\sigma + R\sigma_y\beta)^2 + R^2(\sigma_y\sqrt{1-\beta^2})^2$, $R = I_0/V$, $I_t = By_t$, $B = \frac{(R(L)-L)}{r}$, $\lambda = \sqrt{2r+b^2}$, $a = (1/\sigma_A)(ln(A_B/A))$, $\xi = (1/\sigma_A)(b+\lambda)$, and $b = (1/\sigma_A)(-\sigma_A^2/2)$.

Proof. See in the Appendix. □

The following Table VI summarizes the results when the cash policy is determined according to the System (15), and the debt policy is chosen such that the ex-ante value of equity is maximized. The cash ratio which maximizes shareholder wealth (L^{SH}/V) is positively related to the correlation coefficient β between the rate of occurrence of investment opportunities and the asset value. The reason evolves from the observation that cash can either be used by management to invest, or to defer default during economic distress. In case of a negative correlation, equityholders anticipate that cash holdings are likely to be used to defer default instead of being invested when investment opportunities occur at a higher rate, and tend to be invested in times when opportunities occur at a lower rate. As a consequence, they optimally leave less cash within the firm than in the case of a positive correlation.

INSERT TABLE VI HERE

This finding contrasts the results in Acharya, Almeida, and Campello (2007), who argue that issuing risky debt today and keeping the proceeds in the cash account is equivalent to transferring resources from future states with high cash flows into future states with low cash flows. As a consequence, financially constrained firms with a negative correlation between cash flows and investment opportunities should prefer to issue more debt and to hold higher amounts of cash as a hedging device. The benefit of carrying cash balances is to enhance firm value by relaxing financial constraints in future bad times when valuable investment opportunities tend to arise. My results show that when incorporating that managers act in their own interest, cash is not a suitable mean to transfer resources for investment from good to bad states. The reason is that managers optimally use these funds to avoid default by servicing debt instead of investing during economic distress. Note that it is generally not optimal for equityholders to give debtholders priority over cash reserves during bad times due to its negative impact on future investment. In reality, however, managers' control over cash balances, and common debt covenants restricting investment when the firm struggles induce such a priority.[23] Due to the agency conflict, my model predicts that equityholders optimally leave less cash in firms with high hedging needs because they are more affected by future underinvestment. In fact, Acharya, Almeida, and Campello (2007) find that cash levels and hedging needs are negatively related across both constrained and unconstrained firms (see their Table 2).

Table VI also shows the excess cash and the firm cash ratio. The marginal impact of excess cash on the value of managers' fixed income stream is hardly affected by the correlation between the occurrence of investment opportunities and the asset value. Intuitively, the ability of managers to defer default with cash does not depend on β. Hence, managers reduce their target cash level to a weaker extent than equityholders do when β decreases, because they are only partially compensated with equity. As a consequence, excess cash is higher for lower β.

Finally, the target net debt ratio ($\frac{D(V,L^*)-L^*}{V}$) in the last column of Table VI strongly decreases when the correlation coefficient goes up. The reasons are that the total asset value is more volatile for higher β which induces lower optimal leverage, and that the target corporate cash ratio increases at the same time. Bates, Kahle, and Stulz (2009) report a pronounced decrease in net debt alongside with the increase in corporate cash holdings. Iskandar-Datta and Jia (2010) confirm this finding with international data. My model suggests that an increasing correlation of the occurrence of investment opportunities to the asset value can provide a theoretical explanation.

[23] The availability of cash balances for debt repayments instead of investment in bad times enables the firm to borrow more at initiation. The increased debt capacity does, however, not create firm value from a hedging perspective, since it is equivalent to raising more financing, carrying it as cash balances, and merely paying cash balances back to creditors.

6. Conclusion

This paper incorporates the corporate cash policy into a standard trade-off model of capital structure. Cash can be valuable to capture investment opportunities. The control over cash, however, allows managers to use it to defer, or avoid, default in case of economic distress. As they obtain a fixed income and utility from investments besides the equity linked compensation, the function of cash as a buffer against bankruptcy in hard times induces managers to target excess liquidity. It is shown that the impact of cash holdings on credit risk drives this self-preservation motive for cash. This motive, in turn, explains the relationship of excess cash to firm or industry characteristics, and matches empirically observed properties of cash. Moreover, the analysis quantifies the costs of managers' tendency to overinvest in liquid assets.

The presented results also suggest directives for empirical research. In particular, the relation between cash holdings and credit risk should depend on firm characteristics such as distress costs, or the yield on liquid securities. Additionally, I show that excess cash drives the positive relationship between firms' cash holdings and asset risk. Hence, this dependence should be more pronounced for corporations with lower takeover threats, a more entrenched management, or a higher fixed income of managers.

The model can be extended in many dimensions. A promising path could be to incorporate managers' flexibility to sell assets, and to analyze its impact on the relation between cash holdings and credit risk. Another interesting feature is (costly) renegotiation of both the debt policy and the cash policy.

References

Acharya, Viral, Sreedhar T. Bharath, and Anand Srinivasan, 2007, Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries, *Journal of Financial Economics* 85, 787–821.

Acharya, Viral V., Heitor Almeida, and Murillo Campello, 2007, Is cash negative deb: A hedging perspective on corporate financial policies, *Journal of Financial Intermediation* 16, 515–554.

Acharya, Viral V., Sergei A. Davydenko, and Ilya A. Strebulaev, 2011, Cash holdings and credit risk, *Mimeo.*

Almeida, Heitor, Murillo Campello, and Michael S. Weisbach, 2004, The cash flow sensitivity of cash, *Journal of Finance* 59, 1777–1804.

Altman, Edward I., 1968, Financial ratios, discriminant analysis and the prediction of coprorate bankruptcy, *Journal of Finance* 23, 589–609.

Anderson, Ronald C., and David M. Reeb, 2003, Founding-family ownership and firm performance: Evidence from the S&P 500, *Journal of Finance* 58, 1301–1327.

Anderson, Ronald W., and Suresh Sundaresan, 1996, The design and valuation of debt contracts, *Review of Financial Studies* 9, 37–68.

Asquith, Paul, Robert Gertner, and David Scharfstein, 1994, Anatomy of financial distress: An examination of junk bond issuers, *Quarterly Journal of Economics* 109, 625–658.

Baker, Geroge P., and Brian J. Hall, 2004, CEO incentives and firm size, *Journal of Labor Economics* 22, 767–798.

Baker, Malcolm, and Jeffrey Wurgler, 2002, Market timing and capital structure, *Journal of Finance* 57, 1–32.

Baskin, Jonathan B., 1987, Coprorate liquidity in games of monopoly power, *Review of Economics and Statistics* 69, 312–319.

Bates, Thomas W, Kathleen M. Kahle, and René M. Stulz, 2009, Why do U.S. firms hold so much more cash than they used to?, *Journal of Finance* 64, 1985–2021.

Bebchuk, Lucian, and Yaniv Grinstein, 2005, The growth of executive pay, *Oxford Review of Economic Policy* 21, 283–303.

Begley, Joy J, and Susan Watts, 1996, Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models, *Review of Accounting Studies* 1, 267–284.

Bhamra, Harjoat S., Lars-Alexande Kuehn, and Ilya A. Strebulaev, 2010, The levered equity risk premium and credit spreads: A unified framework, *Review of Financial Studies* 23, 645–703.

Bhattacharya, Sudipto, and Jay R. Ritter, 1983, Innovation and communication: Signalling with partial disclosure, *Review of Economic Studies* 50, 331–346.

Bolton, Patrick, Hui Chen, and Neng Wang, 2009, A unified theory of Tobin's q, corporate investment, financing, and risk management, *NBER working paper No. 214845.*

Bradley, Michael, and Michael R. Roberts, 2004, The structure and pricing of corporate debt covenants, *Mimeo.*

Campello, Murillo, John R. Graham, and Harvey R. Campbell, 2010, The real effect of financial constraints: Evidence from a financial crisis, *Journal of Financial Economics* 97, 470–487.

Chesney, Marc, Monique Jeanblanc-Picqué, and Marc Yor, 1997, Brownian excursions and parisian barrier options, *Advances in Applied Probability* 29, 165–184.

Davydenko, Sergei A., 2007, When do firms default? A study of the default boundary, *Mimeo.*

——— , and Ilya A. Strebulaev, 2007, Strategic actions and credit spreads: An empirical investigation, *Journal of Finance* 62, 2633–2671.

Décamps, Jean-Paul, and Stéphane Villeneuve, 2007, Optimal dividend policy and growth option, *Finance and Stochastics* 11, 3–27.

DeAngelo, Harry, and Linda DeAngelo, 1990, Dividend policy and financial distress: An empirical investigation of troubled NYSE firms, *Journal of Finance* 45, 1415–1431.

Denis, David J., and Valeriy Sibilkov, 2010, Financial constraints, investment, and the value of cash holdings, *Review of Financial Studies* 23, 247–269.

Dittmar, Amy, and Jan Mahrt-Smith, 2007, Corporate governance and the value of cash holdings, *Journal of Financial Economics* 83, 599–634.

——— , and Henri Servaes, 2003, International coprorate governance and corporate cash holdings, *Journal of Financial and Quantitative Analysis* 38, 111–133.

Duffie, Darrell, and David Lando, 2001, Term structures of credit spreads with incomplete accounting information, *Econometrica* 69, 633–664.

Dumas, Bernar, 1991, Super contact and related optimality conditions, *Journal of Economic Dynamics and Control* 15, 675–685.

Faleye, Olubunmi, 2004, Cash and corporate control, *Journal of Finance* 59, 2041–2060.

Fan, Hua, and Suresh M. Sundaresan, 2000, Debt valuation, renegotiation, and optimal dividend policy, *The Review of Financial Studies* 13, 1057–1099.

Faulkender, Michael, and A. Petersen, Mitchell, 2006, Does the source of capital affect capital structure?, *Review of Financial Studies* 19, 45–79.

Fee, Edward C., and Charles J. Hadlock, 2004, Managment turnover across the corporate hierarchy, *Journal of Accounting and Economics* 37, 3–38.

François, Pascal, and Erwan Morellec, 2004, Capital structure and asset prices: Some effects of bankruptcy procedures, *Journal of Business* 77, 387–412.

Froot, Kenneth A., 1992, Intel corporation, *Case Study Harvard Business School, Cambridge, MA.*

Gamba, Andrea, and Alexander Triantis, 2008, The value of financiel flexibility, *Journal of Finance* 63, 2263–2296.

Graham, John R., 2000, How big are the tax benefits of debt?, *Journal of Finance* 53, 1901–1941.

Hackbarth, Dirk, Jianjun Miao, and Erwan Morellec, 2006, Capital structure, credit risk, and macroeconomic conditions, *Journal of Financial Economics* 82, 519–550.

Hall, Brian J., and Kevin J. Murphy, 2002, Stock options for undiversified executives, *Journal of Accounting and Economics* 33, 3–42.

Harford, Jarrad, Sattar A. Mansi, and William F. Maxwell, 2008, Corporate governance and firm cash holdings in the US, *Journal of Financial Economics* 87, 535–555.

Harford, Jarrad, Wayne H. Mikkelson, and Megan Partch, 2003, The effect of cash reserves on corporate investment and performance in industry downturns, *Mimeo.*

Haushalter, David, Sandy Klasa, and William F. Maxwell, 2007, The influence of product market dynamics on a firm's cash holdings and hedging behavior, *Journal of Financial Economics* 84, 797–825.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Hillegeist, Stephen A., Elizabeth K. Keatin, Donald P. Cram, and Kyle G. Lundstedt, 2004, Assessing the probability of bankruptcy, *Review of Accounting Studies* 9, 5–34.

Himmelberg, Charles P., Glenn R . Hubbard, and Darius Palia, 1999, Understanding the determinants of managerial ownership and the link between ownership and performance, *Journal of Financial Economics* 53, 353–384.

Holmstroem, Bengt, and Jean Tirole, 2000, Money, credit, and banking lecture, *Journal of Money, Credit, and Banking* 32, 295–319.

Huberman, Gur, 1984, External financing and liquidity, *Journal of Finance* 39, 895–908.

Hugonnier, Julien N., Semyon Malamud, and Erwan Morellec, 2010, Capital supply uncertainty, cash holdings, and investment, *Mimeo.*

Iskandar-Datta, Mai, and Yonghong Jia, 2010, Why do firms hold so much cash? The international evidence, *Mimeo.*

Jensen, Michael C., 1986, Agency costs of free cash flow, corporate finance, and takeovers, *American Economic Review* 76, 323–329.

Kahan, Marcel, and David Yermack, 1998, Investment opporunities and the design of debt securities, *The Journal of Law, Economics, and Organization* 14, 136–151.

Kalcheva, Ivalina, and Karl V. Lins, 2007, International evidence on cash holdings and expected managerial agency problems, *Review of Financial Studies* 20, 1087–1112.

Keynes, John M, 1936, The general theory of employment, *Interest and Money.*

Kim, Chang-Soo, David C. Mauer, and Ann E. Sherman, 1998, The determinants of coprorate liquidity: Theory and evidence, *Journal of Financial and Quantitative Analysis* 33, 305–334.

Leland, Hayne E., 1994, Corporate debt value, bond covenants, and optimal capital structure, *Journal of Finance* 49, 1213–1252.

Lins, Karl V., Henri Servaes, and Peter Tufano, 2009, What drives corporate liquidity: An international survey of cash hodings and lines of credit, *Mimeo.*

Longstaff, Francis A., and Eduardo S. Schwartz, 1995, A simple approach to valuing risky fixed and floating rate debt, *Journal of Finance* 50, 789–819.

Loughran, Tim, and Jay R. Ritter, 1997, The operating performance of firms conducting seasoned equity offerings, *Journal of Finance* 52, 1823–1850.

Mahmudi, Hamed, and Michael Pavlin, 2010, Corporate payout policy, cash savings, and the cost of consistency: Evidence from a structural estimation, *Mimeo.*

Marchica, Marie-Teresa, and Roberto Mura, 2008, Market frictions and ability to invest: A cash holdings policy perspective, *Mimeo.*

Mauer, David C., and Yixin Liu, 2011, Corporate cash holdings and CEO compensation incentives, *Journal of Financial Economics* 102, 183–198.

Meeks, Geoffrey, and Geoff J. Meeks, 2001, The loser's curse: Accounting for the transaction costs of takeover and the doistortion of takeover motives, *Abacus* 37, 389–400.

Mikkelson, Wayne H., and Megan Partch, 2003, Do persistent large cash reserves hinder performance?, *Journal of Financial and Quantitative Analysis* 38, 275–294.

Miller, Merton H., and Daniel Orr, 1966, A model of the demand for money by firms, *The Quarterly Journal of Economics* 80, 413–435.

Morellec, Erwan, and Boris Nikolov, 2009, Cash holdings and competition, *Mimeo.*

——— , and Norman Schürhoff, 2009, Dynamic capital structure under managerial entrenchment: Evidence from a structural estimation, *Mimeo.*

Murphy, Kevin, 1999, *Executive compensation, in Handbook of Labor Economics* (O. Ashenfelter and D. Card. Elsevier, Amsterdam).

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.

——— , and S. Majluf Nicolàs, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187–221.

Nikolov, Boris, and Toni M. Whited, 2009, Agency conflicts and cash: Estimates from a structural model, *Mimeo.*

Opler, Tim, Lee Pinkowitz, René Stulz, and Rohan Williamson, 1999, The determinants and implications of corporate cash holdings, *Journal of Financial Economics* 52, 3–46.

Ozkan, Aydin, and Neslihan Ozkan, 2004, Corporate cash holdings: An empirical investigation of UK companies, *Journal of Banking and Finance* 28, 2103–2134.

Pinkowitz, Lee, Rene M. Stulz, and R. Williamson, 2003, Do firms in countries with poor protection of investor rights hold more cash?, *Mimeo.*

Pinkowitz, Lee, René M. Stulz, and Rohan Williamson, 2006, Does the contribution of corporate cash holdings and dividends to firm value depend on governance? A cross-country analysis, *Journal of Finance* 61, 2725–2751.

Pinkowitz, Lee F., and Rohan G. Williamson, 2002, What is a dolar worth? The market value of cash holdings, *Mimeo.*

Riddick, Leigh A., and Toni M. Whited, 2009, The corporate propensity to save, *Journal of Finance* 64, 1729–1766.

Shumway, Tyler, 2001, Forecasting bankruptcy more accurately: A simple hazard model, *Journal of Business* 74, 101–124.

Smith, Clifford W., and B. Warner, Jerold, 1979, On financial contracting, *Journal of Financial Economics* 7, 117–161.

Strebulaev, Ilya A., 2007, Do tests of capital structure theory mean what they say?, *Journal of Finance* 62, 1747–1787.

Trigeorgis, Lenos, 1993, The nature of option interactions and the valuation of investments with multiple real options, *Journal of Financial and Quantitative Analysis* 28, 1–20.

Webb, David C., 1987, The importance of incomplete information in explaining the existence of costly bankruptcy, *Econometrica* 54, 279–288.

Yeo, Khim T., and Qui Fasheng, 2003, The value of management flexibility - a real option approach to investment evaluation, *International Journal of Project Management* 21, 243–250.

Yun, Hayong, 2009, The choice of corporate liquidity and corporate governance, *Review of Financial Studies* 22, 1447–1475.

Zmijewski, Mark E., 1984, Methodological issues related to the estimation of financial distress predition models, *Journal of Accounting Research* pp. 59–86.

7. Figures

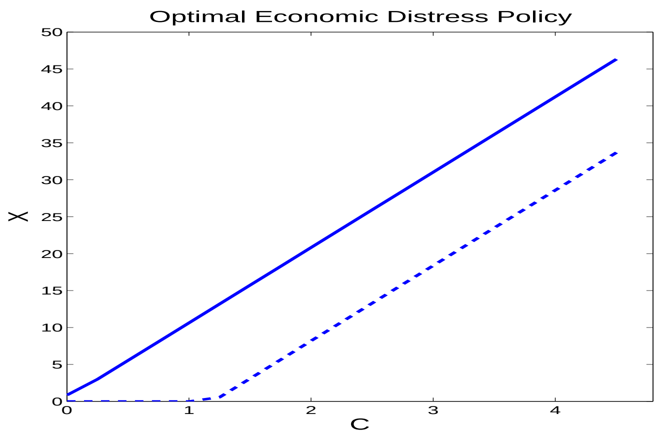

Figure 1. Optimal distress policy. The dashed line shows the threshold where equityholders optimally stop contributing funds in a firm without cash holdings. The solid line is the one of a firm with a cash ratio of $L/V = 0.1$. Standard input parameter values are used.

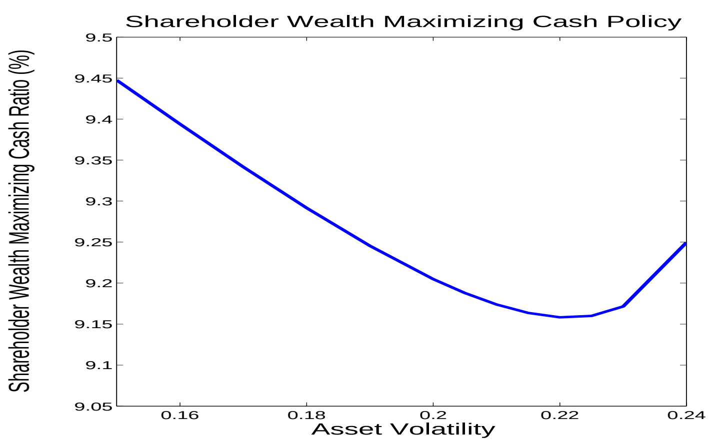

Figure 2. Relationship between asset risk and the shareholder value maximizing cash ratio (L^{SH}/V). The solid line plots optimal cash ratios against the asset volatility. Standard input parameter values are used.

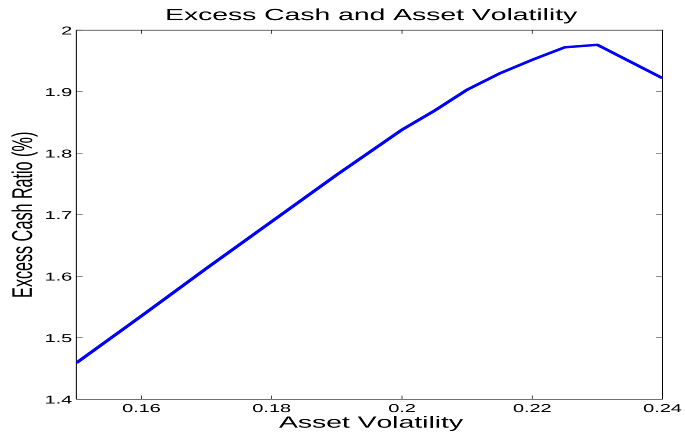

Figure 3. Relationship between asset risk and the excess cash ratio ($(L^* - L^{SH})/V$). The solid line plots excess cash ratios against the asset volatility. Standard input parameter values are used.

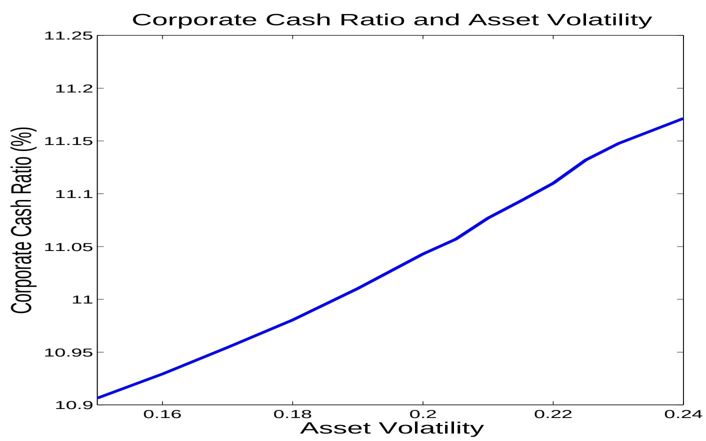

Figure 4. Relationship between asset risk and the corporate cash ratio (L^*/V). The solid line plots corporate cash holdings against the asset volatility. Standard input parameter values are used.

8. Tables

Table I
Cash Holdings and Credit Risk

This table explores the marginal impact of cash on credit risk. The baseline parameters are $V = 100$, $r = 0.06$, $\tau = 0.15$, $\alpha = 0.6$, $\rho = 0.03$, $\delta = 0.05$, $\sigma = 0.2$, $l = 0.02$, $T = 0.05\%$, $A = 4.2$, $\gamma = 0.7$, $y = 0.1$. V is the initial value of the firm's assets, r the riskless interest rate, τ the net tax advantage of debt, α the recovery rate, ρ the economic distress costs, δ the payout ratio, σ the asset volatility, and l the average yield of cash and very liquid securities. A and γ are the parameters of the profit function, and y is the Poisson intensity of the occurrence of investments. Columns 2 to 4 show the direct, indirect and total marginal effects of cash on credit spreads expressed in basis points. The direct effect is the decrease of the credit spread if one additional unit of cash is added, given that the default boundary is left constant. The indirect effect shows the increase of the credit spread caused by the fact that equityholders optimally increase the default threshold with an additional unit of cash. Column 5 shows the total marginal effect of cash on the value of a fixed income stream of 1.

Input Parameter Values	Direct Marginal Effect	Indirect Marginal Effect	Total Marginal Effect	Total Marginal Effect on Value of Fixed Income
Baseline Parameters	-2.8	1.63	-1.17	0.047
$l = 0.01$	-2.47	1.79	-0.68	0.041
$\rho = 0.02$	-3.04	2.24	-0.8	0.044
$\tau = 0.14$	-2.83	1.7	-1.13	0.046
$\delta = 0.04$	-2.3	1.23	-1.07	0.052
$\sigma = 0.21$	-3.09	1.85	-1.24	0.05

Table II
The Impact of Managerial Control on Excess Cash Ratios

This table summarizes the main predictions for excess cash ratios. The target cash level of managers is given by L^*, the one which maximizes shareholder wealth by L^{SH}. l is the average yield of cash and very liquid securities, ρ are the economic distress costs a firm incurs during the time where equityholders stop contributing funds, τ denotes the tax rate, and δ is the payout ratio. Cash Ratios are expressed in percentages of the asset value.

Input Parameter Values	Shareholder Wealth Maximizing Cash Ratio $\frac{L^{SH}}{V}$	Excess Cash Ratio $\frac{L^* - L^{SH}}{V}$	Corporate Cash Ratio $\frac{L^*}{V}$
Baseline Parameters	9.22%	1.82%	11.04%
0% Equity Compensation	9.37%	17.09%	26.46%
30% Equity Compensation	9.47%	4.73%	14.20%
90% Equity Compensation	9.20%	0.32%	9.52%
10% Utility From Investing	9.29%	7.46%	16.75%
20% Utility From Investing	9.27%	12.36%	21.63%
$l = 0.03$	11.88%	2.20%	14.08%
$l = 0.01$	7.34%	1.44%	8.78%
$\rho = 0.04$	9.12%	1.92%	11.04%
$\rho = 0.02$	9.32%	1.69%	11.01%
$\tau = 0.16$	9.15%	1.88%	11.03%
$\tau = 0.14$	9.29%	1.76%	11.05%
$\delta = 0.06$	9.24%	1.43%	10.67%
$\delta = 0.04$	9.25%	2.16%	11.41%

Table III
Excess Cash and the Market for Corporate Control

This table summarizes the impact of the market for corporate control on excess cash. The excess cash ratio and the firm cash ratio are defined as in Table II. T denotes takeover costs, and y^A expresses the ability of the alternative management. Standard input parameter values are used with 10% private utility from investment.

	Excess Cash Ratio	Firm Cash Ratio
Baseline Parameters (T=5%)	7.46%	16.75%
Panel A: $y^A = y = 0.1$		
$T = 0.1\%$	3.64%	12.89%
$T = 0.5\%$	5.86%	15.14%
$T = 0.9\%$	7.46%	16.75%
Panel B: $T = 0.5\%$		
$y^A = 0.102$	3.89%	13.09%
$y^A = 0.1$	5.86%	15.14%
$y^A = 0.098$	7.42%	16.62%

Table IV
Valuation Properties of Cash

This table shows results from the simulated data panel of the true cross section of S&P 500 firms. Panel A and B report averages and quantiles of the market to book values and the marginal values of cash, respectively (absolute values). Panel C shows the average and quantiles of the coefficient on leverage in OLS regressions with fixed effects on the marginal value of cash.

	Coefficient
Panel A: Market to Book Value of Cash	
Mean	1.2959
Median	1.2961
25% Quantile	1.2914
75% Quantile	1.3005
Panel B: Marginal Value of Cash	
Mean	0.6938
Median	0.6937
25% Quantile	0.6924
75% Quantile	0.6951
Panel C: Relation between Leverage and Marginal Value of Cash	
Mean	−0.1176
Median	−0.1170
25% Quantile	−0.1234
75% Quantile	−0.1117

Table V
Determinants of Credit Risk

This table shows the average and quantiles of the coefficients from OLS regressions of credit spreads on explanatory variables. Simulated data panels of the true cross section of S&P 500 firms is used for the regressions.

	Cash Ratio	Leverage	Asset Volatility	Constant
Mean	−1.65	5.23	14.24	−3.62
Median	−1.68	5.22	14.20	−3.60
25% Quantile	−1.87	5.08	13.33	−3.82
75% Quantile	−1.47	5.37	15.24	−3.42

Table VI
Cash Holdings and Correlated Investment Opportunities

This table explores the impact of correlation between the occurrence of investment opportunities and the asset value on cash holdings. σ_y is the volatility of the Poisson intensity of the occurrence of investments, and β denotes the correlation of the intensity to the asset value. Standard input parameter values are used. The shareholder wealth maximizing cash ratio, the excess cash ratio, and the firm cash ratio are defined as in Table II.

	Shareholder Wealth Maximizing Cash Ratio $\frac{L^{SH}}{V}$	Excess Cash Ratio $\frac{L^{*}-L^{SH}}{V}$	Firm Cash Ratio $\frac{L^{*}}{V}$	Net Debt Ratio $\frac{D(V,L^{*})-L^{*}}{V}$
$\sigma_y = 0.2, \beta = 0.5$	10.50%	0.98%	11.48%	25.37%
$\sigma_y = 0.2,\ \beta = 0$	10.28%	1.02%	11.30%	30.36%
$\sigma_y = 0.2,\ \beta = -0.5$	9.96%	1.11%	11.07%	36.35%
$\sigma_y = 0.3, \beta = 0.5$	10.65%	0.95%	11.60%	21.83%
$\sigma_y = 0.3,\ \beta = 0$	10.39%	0.97%	11.36%	28.62%
$\sigma_y = 0.3,\ \beta = -0.5$	9.92%	1.10%	11.02%	37.41%

A. Appendix

Proof of Proposition 1. Firm value if given by

$$\begin{aligned}F(V,L) &= \mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}[\delta V_u + \tau c 1_{V_u>V_B} - \rho V_u 1_{V_u<V_B}]du\big)\\ &+\mathbb{E}_Q\big(\sum_{i=1}^{N_{\theta^{V_B}}} e^{-rt_i}(R_{t_i}(L)-L)1_{V_u>V_B}\big) + lL\mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}1_{V_u>V_B}du\big)\\ &-(1-\alpha)\mathbb{E}_Q(e^{-r\theta^{V_B}}V_{\theta^{V_B}}) + L\mathbb{E}_Q(e^{-r\theta^{V_B}}).\end{aligned} \tag{1}$$

Following the proof of François and Morellec (2004) on page 406 − 408, it is straightforward to show that $\mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}[\delta V_u + \tau c 1_{V_u>V_B} - \rho V_u 1_{V_u<V_B}]du\big) - (1-\alpha)\mathbb{E}_Q(e^{-r\theta^{V_B}}V_{\theta^{V_B}}) =$

$$\begin{aligned}&V - \frac{\delta V_B}{\lambda^2+(\sigma+b)\lambda}(\frac{V}{V_B})^{-\xi} - \delta\frac{V_B}{\lambda}(\frac{V}{V_B})^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma-b}\\ &+\frac{\tau c}{r} - \frac{\tau c}{r}(\frac{\lambda-b}{2\lambda} + \frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{V}{V_B})^{-\xi}\\ &-(1-\alpha)V_B(\frac{V}{V_B})^{-\xi}\frac{\Phi(-(b+\sigma)\sqrt{d})}{\Phi(\lambda\sqrt{d})}\\ &+(\delta-\rho)\big((1/\delta)\big(V - \frac{\phi(-(\sigma+b)\sqrt{d})}{\phi(\lambda\sqrt{d})}V_B(\frac{V}{V_B})^{-\xi}\big)\\ &-\big(V/\delta - \frac{V_B}{\lambda^2+(\sigma+b)\lambda}(\frac{V}{V_B})^{-\xi} - \frac{V_B}{\lambda}(\frac{V}{V_B})^{-\xi}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})}\frac{1}{\lambda-\sigma-b}\big)\big).\end{aligned} \tag{2}$$

To derive

$$\mathbb{E}_Q\big(\sum_{i=1}^{N_{\theta^{V_B}}} e^{-rt_i}(R_{t_i}(L)-L)1_{V_u>V_B}\big) + lL\mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}1_{V_u>V_B}du\big), \tag{3}$$

I first express the expected instantaneous return from cash as

$$\Gamma(L) \overset{!}{=} y(R(L)-L) + lL \tag{4}$$

by using the compensator of a compound Poisson process. If the Poisson intensity, the profit function, and the Brownian motion $(W_t)_t$ are independent, one can rewrite Expression (3) as

$$\Gamma(L)\mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}1_{V_u>V_B}du\big). \tag{5}$$

Following the steps in François and Morellec (2004), the solution to Expression (5) is given by

$$\frac{\Gamma(L)}{r} - \frac{\Gamma(L)}{r}(\frac{\lambda-b}{2\lambda} + \frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{V}{V_B})^{-\xi}. \tag{6}$$

Finally, the Laplace transform of θ^{V_B}, i.e., $\mathbb{E}(e^{-\frac{x^2}{2}\theta^{V_B}})$, is given by $\frac{e^{ax}}{\Phi(x\sqrt{d})}$ (Chesney, Jeanblanc-Picqué, and Yor 1997). Hence, $L\mathbb{E}_Q(e^{-r\theta^{V_B}})= L\frac{e^{a\sqrt{2r}}}{\Phi(\sqrt{2rd})}$. Combining the individual results and rearranging yields the expression. □

Proof of Proposition 2. The first part of the proposition, Equation (9), directly follows from the proof of Proposition 1. The second part is obtained by replacing V by V_B in Equation (9), and comparing the result to $\alpha V_B + L$. □

Proof of Proposition 3. The value of the fixed income stream to managers is given by $\mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}\omega du\big) = \mathbb{E}_Q\big(\int_0^{\infty} e^{-ru}\omega du\big) - \mathbb{E}_Q\big(\int_{\theta^{V_B}}^{\infty} e^{-ru}\omega du\big)$. The first term on the right hand side is simply equal to $\frac{\omega}{r}$, the second term corresponds to the value of an infinite stream obtained at θ^{V_B}. As the Laplace transform of θ^{V_B} is know, this second term can be expressed as $\frac{\omega}{r}\frac{e^{a\sqrt{2r}}}{\Phi(\sqrt{2rd})}$, which, subtracted from $\frac{\omega}{r}$, gives the desired result.

The value of the equity share to managers corresponds to Ψ times the continuation value of equity minus the raised cash.

To derive the private utility from investing cash, I use the compensator of a compound Poisson process to express the expected instantaneous utility as $y\eta log(L)$. Hence, $y\eta log(L)\mathbb{E}_Q\big(\int_0^{\theta^{V_B}} e^{-ru}1_{V_u>V_B}du\big) = \frac{y\eta ln(L)}{r}\big(1-(\frac{\lambda-b}{2\lambda}+\frac{\lambda+b}{2\lambda}\frac{\Phi(-\lambda\sqrt{d})}{\Phi(\lambda\sqrt{d})})(\frac{V}{V_B})^{-\xi}\big)$. □

Proof of Proposition 4 The intensity follows

$$dy_t = \sigma_y y_t dW_t^y, \tag{7}$$

with $E(dW_t dW_t^y) = \beta dt$. One can write the following decomposition

$$dy_t = \sigma_y y_t(\beta dW_t + \sqrt{1-\beta^2}dW_t^{\epsilon}), \tag{8}$$

where W_t^{ϵ} is a Brownian motion independent of W_t. Define the stochastic part of the total asset value (invested assets plus investment opportunities) as $A_t = V_t + E_Q\int_t^{\infty}(R(L)-L)y_u e^{-ru}du = V_t + \frac{(R(L)-L)}{r}y_t$, and the ex-post default strategy, A_B, in terms of A_t.

The value of the investment opportunities follows

$$dI_t = \sigma_y B y_t(\beta dW_t + \sqrt{1-\beta^2}dW_t^{\epsilon}), \tag{9}$$

where $B = \frac{(R(L)-L)}{r}$. For simplicity, assume that $\delta = r$ such that the total payout is equal to $r_t A_t$, and that the recovery rate and distress costs are defined on the stochastic asset value. The dynamics of A_t can now be written as

$$dA_t = \sigma_A A_t dW_t^A, \tag{10}$$

where W^A is a standard Brownian motion, $\sigma_A^2 = (\sigma + R\sigma_y\beta)^2 + R^2(\sigma_y\sqrt{1-\beta^2})^2$, $R = I_0/V$, and $I_0 = By_0$.

Once the evolution of the stochastic asset value is known, we can calculate firm value as

$$\begin{aligned} F(V,L,y) = \mathbb{E}_Q\big(\int_0^{\theta^{A_B}} & e^{-ru}[rA_u + \tau c 1_{A_u > A_B} - \rho A_u 1_{A_u < A_B}]du\big) \\ & + lL\mathbb{E}_Q\big(\int_0^{\theta^{A_B}} e^{-ru} 1_{A_u > A_B} du\big) - (1-\alpha)\mathbb{E}_Q(e^{-r\theta^{A_B}} A_{\theta^{A_B}}) + L\mathbb{E}_Q(e^{-r\theta^{A_B}}), \end{aligned} \tag{11}$$

and debt value as

$$D(V,L,y) = E_Q(\int_0^{\theta^{A_B}} e^{-ru} c 1_{A_u > A_B} du) + \alpha E_Q(e^{-r\theta^{A_B}} A_{\theta^{A_B}}) + L\mathbb{E}_Q(e^{-r\theta^{A_B}}). \tag{12}$$

Both expressions can be solved as in the previous proofs. □

PART III: CURRICULUM VITAE

Marc Arnold - Curriculum Vitae — November 2011

Contact Information

University of Zurich
Department of Banking and Finance
Plattenstrasse 14,
CH-8032 Zurich, Switzerland

Office: +41 44 634 39 57
Mobile: +41 79 756 69 48
E-mail: marc.arnold@bf.uzh.ch

Citizenship: Swiss — Date of Birth: 21.07.1977

Areas of Interests

Research: Theoretical Corporate Finance, Credit Risk, Financial Intermediation

Education

University of Zurich, Switzerland

Swiss Finance Institute PhD Program in Finance (summa cum laude), Doctoral Studies — **2007 – 2011**

- Advisor: Professor Alexander F. Wagner, PhD.
- Working Title: Essays on Credit Risk

Lic. oec. publ. (summa cum laude), Master Degree — **1998 – 2003**

- Focus: Finance
- Master Thesis: Mögliche Behavioral Faktoren in der Arbitrage Pricing Theory
- Bachelor Thesis: Eignung der Arbitrage Pricing Theory als Grundlage für das Portfoliomanagement

Working Papers

- The Impact of Managerial Control over Cash on Credit Risk and Financial Policy, 2011
- Macroeconomic Conditions, Growth Opportunities and the Cross Section of Credit Risk, *Swiss Finance Institute Research Paper No. 10-19*, 2011 with Alexander F. Wagner and Ramona Westermann, invitation to revise and resubmit to the Journal of Financial Economics.
- Private Information and Callable Credit Default Swaps, 2011

Conference and Seminar Speeches

- Research Seminar at University of Luxembourg, Luxembourg (November 2011)
- Research Seminar at University of Vienna, Austria (September 2011)
- Research Seminar at University of St. Gallen, Switzerland (September 2011)
- C.R.E.D.I.T Conference Scuola Grande San Giovanni Evangelista Venice, Italy (October 2010)
- 9^{th} Swiss Doctoral Workshop in Finance, Gerzensee, Switzerland (June 2010)
- Research Seminar at University of Zurich, Switzerland (May 2010)
- Research Seminar at University of Zurich, Switzerland (April 2010)
- 8^{th} Swiss Doctoral Workshop in Finance, Gerzensee, Switzerland (June 2009)
- Research Seminar at University of Zurich, Switzerland (March 2009)
- 7^{th} Swiss Doctoral Workshop in Finance, Gerzensee, Switzerland (June 2008)

Professional Experience

Department of Banking and Finance - University of Zurich
Research and Teaching Assistant — **2006 – present**

- Research Assistant for Prof. Alexander Wagner, PhD. (since 2006)
- Teaching Assistant for PhD Courses in Corporate Finance (since 2007)
- Master Theses Supervision (since 2007)

CREDIT SUISSE-Zurich

Junior Product Structurer **2003 –2005**

- Structured Derivatives Initiation and Consulting
- Career Start Program
- Project Participation in the Departement Trading and Sales

Awards and Scholarships

ProDoc Grant of the Swiss National Science Foundation **2008 – 2011**

Swiss Finance Institute Scholarship **2007 – 2008**